0 EP 505 387

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 : C07H 21/00, C12N 1/20, 15/00 C07K 13/00 — A1

(11) International Publication Number: WO 91/07423

(43) International Publication Date: 30 May 1991 (30.05.91)

(21) International Application Number: PCT/US90/06015

(22) International Filing Date: 19 October 1990 (19.10.90)

(30) Priority data:
438,775 17 November 1989 (17.11.89) US

(71) Applicant: ARCH DEVELOPMENT CORPORATION [US/US]; 1115-25 East 58th Street, Chicago, IL 60637 (US).

(72) Inventors: LIAO, Shutsung ; 5632 South Woodlawn Avenue, Chicago, IL 60637 (US). CHANG, Chawnshang ; 5644 South Harper Avenue, Chicago, IL 60637 (US).

(74) Agent: MEINERT, M., C.; Marshall, O'Toole et al., Two First National Plaza, Suite 2100, Chicago, IL 60603 (US).

(81) Designated States: AT (European patent), BE (European patent), CH (European patent), DE (European patent), DK (European patent), ES (European patent), FR (European patent), GB (European patent), GR (European patent), IT (European patent), JP, LU (European patent), NL (European patent), SE (European patent).

Published
With international search report.

(54) Title: DNA BINDING PROTEINS INCLUDING ANDROGEN RECEPTOR

(57) Abstract

Disclosed are DNA sequences encoding DNA binding polypeptides including androgen receptor (AR) and TR2 polypeptides. Illustratively, human and rat AR-cDNA have 79 kD and 98 kD polypeptide expression products which are immunoprecipitable by human auto-immune anti-androgens receptor antibodies and are capable of binding androgens specifically and with high affinity. Also disclosed are antibodies and immunological methods and materials for detection of androgen receptor and TR2 polypeptides and hybridization methods and materials for detection of AR-and TR2-related nucleic acids.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	FI	Finland	ML	Mali
AU	Australia	FR	France	MN	Mongolia
BB	Barbados	GA	Gabon	MR	Mauritania
BE	Belgium	GB	United Kingdom	MW	Malawi
BF	Burkina Faso	GN	Guinea	NL	Netherlands
BG	Bulgaria	GR	Greece	NO	Norway
BJ	Benin	HU	Hungary	PL	Poland
BR	Brazil	IT	Italy	RO	Romania
CA	Canada	JP	Japan	SD	Sudan
CF	Central African Republic	KP	Democratic People's Republic of Korea	SE	Sweden
CG	Congo			SN	Senegal
CH	Switzerland	KR	Republic of Korea	SU	Soviet Union
CI	Côte d'Ivoire	LI	Liechtenstein	TD	Chad
CM	Cameroon	LK	Sri Lanka	TG	Togo
DE	Germany	LU	Luxembourg	US	United States of America
DK	Denmark	MC	Monaco		
ES	Spain	MG	Madagascar		

DNA BINDING PROTEINS INCLUDING ANDROGEN RECEPTOR

CROSS-REFERENCE TO RELATED APPLICATIONS

The present invention is a continuation-in-part of co-pending U.S. Patent Application Serial No. 07/312,763, filed February 21, 1989; which in turn is a continuation-in-part of expressly abandoned U.S. Patent Application Serial No. 07/253,807, filed October 5, 1988; which in turn is a continuation-in-part of expressly abandoned U.S. Patent Application Serial No. 06/176,107, filed March 30, 1988

BACKGROUND OF THE INVENTION

The present invention relates generally to DNA binding regulatory proteins and more particularly to DNA sequences encoding androgen receptor protein and novel DNA binding proteins designated TR2, to the polypeptide products of recombinant expression of these DNA sequences, to peptides whose sequences are based on amino acid sequences deduced from these DNA sequences, to antibodies specific for such proteins and peptides, and to procedures for detection and quantification of such proteins and nucleic acids related thereto.

There are five major classes of steroid hormones: progestins, glucocorticoids, mineralocorticoids, androgens, and estrogens. Receptor proteins, each specific for a steroid hormone, are distributed in a tissue specific fashion and in target cells, steroid hormones can form specific complexes with corresponding intracellular receptors. [Jensen, et al., <u>Proc. Nat'l. Acad. Sci. (USA)</u>, <u>59</u>:632 (1968); Gorski, et al., <u>Ann. Rev. Physiol.</u>, <u>38</u>:425-450 (1976); and Liao, et al., page 633 in <u>Biochemistry of Hormones</u>, H.L.J. Makin, ed. (Blackwell Sci. Publ. Oxford, 1984)]. The hormonal regulation of gene expression appears to involve interaction of steroid receptor complexes with certain seg-

ments of genomes and modulation of specific gene transcription. See, e.g., Ringold, Ann. Rev. Pharmacol. Toxicol., 25:529 (1985); and Yamamoto, Ann. Rev. Genet., 19:209 (1985). Many of the primary effects of hormones involve increased transcription of a subset of genes in specific cell types.

The successful cloning of e.g., cDNAs coding for various steroid receptors has allowed the structural and functional analysis of different steroid receptor domains involved in steroid and DNA binding. See, e.g., Hollenberg, et al., Nature (London), 318:635 (1985); Miesfeld, et al., Cell, 46:389 (1986); Danielsen, et al., EMBO J., 5:2513 (1986); Greene, et al., Science, 231:1150 (1986); Green, et al., Nature (London), 320:134 (1986); Krust, et al., EMBO J., 5:891 (1986); Loosfelt, et al., Proc. Nat'l. Acad. Sci. (USA), 83:9045 (1986); Conneely, et al., Science, 233:767 (1987); Law, et al., Proc. Nat'l. Acad. Sci. (USA), 84:2877 (1987); Misrahi, et al., Biochem. Biophys. Res. Commun., 143:740 (1987); Arriza, et al., Science, 237:268 (1987); Sap, et al., Nature (London), 324:635 (1986); Weinberger, et al., Nature (London), 318:641 (1986); Benbrook, et al., Science, 238:788 (1987); and Evans, Science, 240:889 (1988).

Androgens, such as testosterone, are responsible for the development of male secondary sex characteristics and are synthesized primarily in testis. Cloning of a cDNA for androgen receptor (AR) has been difficult because, until recently, monospecific antibodies against AR have not been available for screening cDNA libraries. An abstract by Govindan, et al., J. Endocrinol. Invest., 10 (Suppl. 2) (1987), reported the isolation of cDNA clones encoding human androgen receptor isolated from a human testis λgt-11 cDNA library using synthetic oligonucleotides homologous to human glucocorticoid, estradiol, and progesterone

receptors as probes. The expressed protein reportedly bound tritium-labelled DHT (dihydrotestosterone) with high affinity and specificity. However, no nucleotide or amino acid sequence analysis was provided for full length androgen receptors, nor was any description provided concerning isolation of the full length putative androgen receptor clones.

Recently, Chang, C., et al., Science, 240:324 (April 15, 1988), co-authored by the inventors herein, described cDNAs encoding androgen receptors obtained from human testis and rat ventral prostrate cDNA libraries. These cDNAs for human and rat androgen receptor were reported to be long enough to code for 94 kDa and 76 kDa receptors. The molecular weights were derived with the assistance of a software program known as: DNA Inspector II (Textco West Lebanon, New Hampshire) open reading frame analysis. With a new DNA Inspector IIe program, hAR (918 amino acids) has an estimated molecular weight 98,608 and rAR (902 amino acids) has a molecular weight of 98,133. Therefore, the reported "94 kDa" AR is now termed "98 kDa" AR; and the hAR or rAR polypeptides, from the second ATG/Met, reported as "76 kDa" are now termed "79 kDa". See also, Chang, C., et al., Proc. Nat'l. Acad. Sci. (USA), 85:7211 (October 5, 1988) also co-authored by the inventors herein.

In contrast, L. ahn, D., et al., Science, 240:327 (1988), using libraries from human epididymis and cultured human foreskin fibroblasts obtained a human cDNA which was expressed in monkey kidney (COS) cells to yield a protein, present in the cytosol, capable of binding androgens. This cDNA, however, was only sufficient to code for a receptor having an estimated molecular weight of 41,000. Therefore, the cDNA obtained only coded for a portion of AR.

Of interest to the present invention is Young, et al., Endocrinol., 123:601 (1988), wherein the production of anti-AR monoclonal antibodies was reported. Anti-AR autoantibodies were identified in the sera of prostate cancer patients, as described in Liao, S., et al., Proc. Nat'l. Acad. Sci. (USA), 82:8345 (1984) (one of the co-inventors herein), and were characterized with respect to their titer, affinity, and specificity. Subsequently, lymphocytes from the blood of those patients having high antibody titers were isolated, transformed with Epstein-Barr Virus (EBV), and cloned for anti-AR monoclonal antibody production. These monoclonal antibodies were found to interact with androgen receptors from rat prostate. An attempt to scale-up antibody production resulted in a decline of antibody secretion. It is not uncommon for transformed B-cells to be more unstable than hybridoma cells. Kozbor, et al., Eur. J. Immunol., 14, 23 (1984). Because of the instability associated with such cell lines, an alternate source of monoclonal antibodies is preferred.

There thus exists a need in the art for information concerning the primary structural conformation of androgen receptor protein and other DNA binding proteins such as might be provided by knowledge of human and other mammalian DNA sequences encoding the same. Availability of such DNA sequences would make possible the application of recombinant methods to the large scale production of the proteins in procaryotic and eukaryotic host cells, as well as DNA-DNA, DNA-RNA, and RNA-RNA hybridization procedures for the detection, quantification and/or isolation of nucleic acids associated with the proteins. Possession of androgen receptor and related DNA-binding proteins and/or knowledge of the amino acid sequences of the same would make possible, in turn, the development of monoclonal and

polyclonal antibodies thereto (including antibodies to protein fragments or synthetic peptides modeled thereon) for the use in immunological methods for the detection and quantification of the proteins in fluid and tissue samples as well as for tissue specific delivery of substances such as labels and therapeutic agents to cells expressing the proteins.

BRIEF SUMMARY OF THE INVENTION

The present invention provides novel purified and isolated DNA sequences encoding androgen receptor protein and a structurally related protein, designated TR2 protein, which also has DNA binding (and hence DNA replication or transcription regulatory) capacity. In presently preferred forms, novel DNA sequences comprise cDNA sequences encoding human and rat androgen receptor and human TR2 protein. Alternate DNA forms such as genomic DNA, and DNA prepared by partial or total chemical synthesis from nucleotides as well as DNA with deletions or mutations, is also within the contemplation of the invention.

Association of DNA sequences provided by the invention with homologous or heterologous species expression control DNA sequences, such as promoters, operators, regulators and the like, allows for *in vivo* and *in vitro* transcription to form messenger RNA which, in turn, is susceptible to translation to provide androgen receptor and TR2 proteins, and related poly- and oligo-peptides in large quantities. In a presently preferred DNA expression system of the invention, AR and TR2 encoding DNA is operatively associated with a viral (T7) regulatory (promoter) DNA sequence allowing for *in vitro* transcription and translation in a cell free system to provide, e.g., a 79 kD and 98 kD human androgen receptor (hAR) protein, 79 kD and 98 kD rat

androgen receptor (rAR) protein and smaller forms of these proteins; as well as TR2 protein, including 20 kD, 52 kD, and 67 kD species.

Incorporation of DNA sequences into procaryotic and eucaryotic host cells by standard transformation and transfection processes, potentially involving suitable viral and circular DNA plasmid vectors, is also within the contemplation of the invention and is expected to provide useful proteins in quantities heretofore unavailable from natural sources. Systems provided by the invention included transformed E. coli DH5α cells, deposited January 25, 1989, with the American Type Culture Collection, 12301 Parklawn Drive, Rockville, Maryland 20852 in accordance with the U.S. Patent and Trademark Office's requirements for microorganism deposits, and designated EC-hAR3600 under A.T.C.C. Accession No. 67879; EC-rAR 2830, A.T.C.C. No. 67878; EC-TR2-5, A.T.C.C. No. 67877; and EC TR2-7, A.T.C.C. No. 67876; as well as transformed E. coli DH5α cells, deposited November 14, 1989 and designated EC TR2-11 under A.T.C.C. No. 68173. Use of mammalian host cells is expected to provide for such post-translational modifications (e.g., truncation, glycosylation, and tyrosine, serine or threonine phosphorylation) as may be needed to confer optimal biological activity on recombinant expression products of the invention.

Novel protein products of the invention include polypeptides having the primary structural conformation (i.e., amino acid sequence) of AR and TR2 proteins as well as peptide fragments thereof and synthetic peptides assembled to be duplicative of amino acid sequences thereof. Proteins, protein fragments, and synthetic peptides of the invention are projected to have numerous uses including therapeutic, diagnostic and prognostic uses and will provide the basis for preparation of monoclonal and polyclonal antibodies specifi-

cally immunoreactive with AR and TR2 proteins. Preferred protein fragments and synthetic peptides include those duplicating regions of AR and TR2 proteins which are not involved in DNA binding functions and the most preferred are those which share at least one antigenic epitope with AR and TR2 proteins.

Also provided by the present invention are polyclonal and monoclonal antibodies characterized by their ability to bind with high immunospecificity to AR and TR2 proteins and to their fragments and peptides, recognizing unique epitopes which are not common to other proteins especially DNA binding proteins.

Illustratively provided according to the present invention are monoclonal antibodies, designated AN1-6, AN1-7, AN1-15; and produced by hybridoma cell lines designated H-AN1-6, H-AN1-7, H-AN1-15; deposited January 25, 1989, under Accession Nos. HB 10,000; HB 9,999; and HB 10,001, respectively, with the American Type Culture Collection, 12301 Parklawn Drive, Rockville, Maryland 20852 in accordance with the U.S. Patent and Trademark Office's requirements for microorganism deposits. These antibodies are characterized by (a) capacity to bind androgen receptors from rat ventral prostate and synthetic peptides having sequences predicted from the structure of hAR-cDNA and rAR-cDNA; (b) specific immunological reactivity with, and capacity to reversibly immunobind to, naturally occurring and recombinant androgen receptors, in native and denatured conformations; and (c) specific immunological reactivity with, and capacity to reversibly immunobind to, proteinaceous materials including all or a substantially, immunologically significant, part of an amino acid sequence duplicative of that extant at residues 331 through 577 of hAR and corresponding amino acid sequences in rAR.

Also provided according to the present invention are monoclonal antibodies to TR2 proteins designated A-TR-2-11a. These antibodes are characterized by their capacity to bind TR2 proteins as well as synthetic peptides having sequences predicted from the structure of hTR-2-cDNA.

The monoclonal antibodies of the invention can be used for affinity purification of AR and TR-2 receptor from human or rat prostate, and other sources such as AR-rich organs and cultured cells.

Also provided by the present invention are novel procedures for the detection and/or quantification of normal, abnormal, or mutated forms of AR and TR2, as well as nucleic acids (e.g., DNA and mRNA) associated therewith. Illustratively, antibodies of the invention may be employed in known immunological procedures for quantitative detection of AR and TR2 proteins in fluid and tissue samples, of DNA sequences of the invention (particularly those having sequences encoding DNA binding proteins) that may be suitably labelled and employed for quantitative detection of mRNA encoding these proteins.

Among the multiple aspects of the present invention, therefore, is the provision of (a) novel AR and TR2-encoding DNA sequences set out in Figure 3, as well as (b) AR and TR2-encoding DNA sequences which hybridize thereto under hybridization conditions of the stringency equal to or greater than the conditions described herein and employed in the initial isolation of cDNAs of the invention, and (c) DNA sequences encoding the same allelic variant, or analog AR and TR2 polypeptides through use of, at least in part, degenerate codons. Correspondingly provided are viral or circular plasmid DNA vectors incorporating such DNA sequences and procaryotic and eucaryotic host cells transformed or transfected with such DNA sequences and

vectors as well as novel methods for the recombir c production of AR and TR2 proteins through cultured growth of such hosts and isolation of these proteins from the hosts or their culture media.

Preferred polypeptide products of the invention include the approximately 79 kD (starting from the second ATG/Met) and 98 kD (starting from the first ATG/Met) hAR polypeptides having the deduced amino acid sequence of 734 and 918 residues, respectively, as set out in Figure 3. Also preferred are the 79 kD and 98 kD rAR species polypeptides having the deduced sequence of 733 and 902 residues set out in Figure 3; as well as the 20 kD, 52 kD, and 67 kD species human TR2 polypeptides having the deduced amino acid sequences of 184, 483, 467, and 603 residues set out in Figures 4, 5, and 6, respectively. The preferred 79 kD and 98 kD hAR and rAR polypeptides may be produced <u>in vitro</u> and are characterized by a capacity to specifically bind androgens with high specificity and by their immunoprecipitatability by human auto-immune anti-androgen receptor antibodies. The preferred 20 kD, 52 kD, and 67 kD TR2 polypeptides may be produced <u>in vitro</u> and are characterized by their ability to interact with TR-2 antibodies and to interact with DNA.

Other aspects and advantages of the present invention will be apparent upon consideration of the following detailed description thereof which includes numerous illustrative examples of the practice of the invention, reference being made to the drawing wherein:

Figure 1 illustrates the strategy employed in construction of a human androgen receptor cDNA vector;

Figure 2 illustrates the strategy employed in construction of rat androgen receptor cDNA vectors;

Figure 3 provides a 3715 base pair nucleotide sequence for a human androgen receptor (hAR) DNA clone and the deduced sequence of 734 and 918 amino acid

residues for hAR proteins; and in addition provides a 3218 base pair nucleotide sequence for a rat androgen receptor (rAR) DNA clone and the deduced sequences of 733 and 902 amino acids for two rAR species;

Figure 4 provides a 2029 base pair nucleotide sequence for a human TR2 DNA clone and a deduced sequence of 483 amino acids for a "TR2-5" species with a calculated molecular weight of 52,982 daltons and a deduced sequence of 184 amino acids for a "TR2-7" species with a calculated molecular weight of 20,528 daltons.

Figure 5 provides a 1785 base pair nucleotide sequence for a human TR2 DNA clone and a deduced sequence of 467 amino acids for a "TR2-9" species with a calculated molecular weight of 50,849 daltons; the amino acid sequence in the DNA-binding domain is boxed. The polyadenylation signal AATAAA is underlined.

Figure 6 provides a 2221 base pair nucleotide sequence for a human TR2 DNA clone and a deduced sequence of 603 amino acids for a "TR2-11" species with a calculated molecular weight of 67,223; the amino acid sequence in the DNA-binding domain is boxed. The polyadenylation signal AATAAA is underlined.

Figure 7 provides an amino acid sequence alignment of the cysteine-rich DNA binding domain of human androgen receptor, glucocorticoid receptor, mineralocorticoid receptor, progesterone receptor, estrogen receptor, TR2, rat AR, chick vitamin D receptor (c-VDR), and the v-erb A oncogene product of avian erythroblastosis virus.

Figure 8 provides a schematic comparison of the four variants of human TR2 receptors: TR2-5; TR2-7; TR2-9; and TR2-11; numbers above the boxes indicate the positions of amino acid residues. The DNA-binding domain (DNA) and the hormone-binding domain (Hormone) are shown. The sequences for TR2-5, TR2-9, and TR2-11 are identical from amino acid number 1 to 464.

Figures 9, 10, and 11 illustrate, respectively, the in-frame fusion of three different parts of the AR gene (the N-terminal, the DNA-binding domain and the androgen-binding domain) to the N-terminal half of the trpE gene using pATH expression vectors.

DETAILED DESCRIPTION

The following examples illustrate practice of the invention. Example 1 relates to the isolation, preparation, and partial structural analysis of cDNA for human and rat androgen receptors. Example 2 relates to confirmation of the presence on the human X-chromosome of an AR-type cDNA sequence. Example 3 relates to the preparation of human and rat cDNAs containing AR-type cDNA from different clones and ligation into the pGEM-3Z plasmid. Example 4 relates to transcription and translation of the AR-type cDNA plasmid DNA. Example 5 relates to steroid binding activity of the expression product of Example 4. Example 6 relates to the binding activity of the expression product of Example 4 to human auto-antibodies. Example 7 relates to the characterization of TR2-cDNA. Example 8 relates to the *in vitro* transcription and translation of TR2-cDNA. Example 9 relates to the binding activity of TR2-cDNA expression product. Example 10 provides a schematic comparison of the four variants of human TR2 receptors. Example 11 relate to the androgen regulation of TR2 mRNA levels in the rat ventral prostate. Example 12 relates to recombinant expression systems of the invention. Example 13 relates to the production of fusion proteins and their use in producing polyclonal and monoclonal antibodies according to the invention. Example 14 relates to use of DNA probes of the inventions. Example 15 relates to development of transgenic animals by means of DNA sequences of the invention.

These examples are for illustrative purposes only and are not intended in any way to limit the scope of the invention.

EXAMPLE 1

Preparation and Partial Structural Analysis of cDNA for Human and Rat Androgen Receptors

The isolation of cDNA for human androgen receptor (hAR) and rat androgen receptor (rAR) was accomplished using λGT11 cDNA libraries. The human testis and prostate λGT11 libraries were obtained from Clontech Co., Palo Alto, California and a rat ventral prostate λGT11 library in E. coli Y1090 was constructed as described in Chang, et al., J. Biol. Chem., 262:11901 (1987). In general, clones were differentiated using oligonucleotide probes specific for various steroid receptors. The cDNA libraries were initially screened with a set of 41-bp oligonucleotide probes designed for homology to nucleotide sequences in the DNA-binding domain of glucocorticoid receptors (GR), estrogen receptors (ER), progesterone receptors (PR), mineralocorticoid receptors (MR), and the v-erb A oncogene product of avian erythroblastosis virus. The set of probes had the following sequence: TGTGGAAGCTGT/CAAAGTC/ATTCTTAAAAGG/AGCAA/GTGGAAGG.

The plaques were replicated on a nitrocellulose filter and screened with 5'-end ^{32}P-labeled 41-bp oligonucleotide probes. The conditions of hybridization were 25% formamide, 5X Denhardt's solution (0.1% Ficoll 400, 0.1% polyvinylpyrrolidone, 0.1% bovine serum albumin), 0.1% SDS, 5X SSC (1X SSC is 150 mM NaCl, 15 mM sodium citrate), 100 μg/ml denatured salmon sperm DNA, and 1 μg/ml poly(A) at 30°C. Filters were washed with a solution containing 0.1% SDS, 0.05% sodium pyrophosphate and 0.4X SSC at 37°C.

A less stringent hybridization condition (2X SSC at 37°C) was used for the first screen employing the 41 bp probes. The remaining clones were then probed again at more stringent conditions by reducing the concentration of SSC, eventually to 0.4X SSC at 37°C, or by increasing the temperature, or by increasing the concentration of formamide. In some procedures, 5X SSC, 8% dextran sulfate, and 20% formamide, at 42°C was employed and the result was equivalent to that obtained with 0.6X SSC.

From approximately 3 X 10^6 human testis recombinants and 6 X 10^5 rat ventral prostate recombinants, 302 and 21 positive clones, respectively, were obtained.

Based on the assumption that AR might have a cysteine-rich DNA binding domain highly homologous to the DNA-binding regions of other steroid receptors, positive clones from the first screenings were probed with 5'-end ^{32}P-labeled 24-bp oligonucleotides specific for the various steroid receptors for the possible presence of cDNA for AR through a process of elimination. The GR-cDNA clones were eliminated by screening with two GR-specific 24-bp probes that had nucleotide sequences identical to nucleotide segments immediately next to the 5'-end or the 3'-end of the DNA binding-region of hGR-cDNA , i.e., TGTAAGCTCTCCTCCATCCAGCTC and CAGCAGGCCACTACAGGAGTCTCA. 244 and 14 clones, respectively, were eliminated as hGR- and rGR-cDNA clones.

Using similar procedures involving four 24-bp probes for the 5'-end of PR(CCGGATTCAGAAA/GCCAGT/-CCAGAGC) and two 24-bp probes for the 3'-end of ER(GCA/-CGACCAGATGGTCAGTGCCTTG), no ER- or PR-cDNA clones were detected in the human testis library. In the rat prostate library, no ER-cDNA clones were detected but one positive clone was obtained with hPR-specific 24 bp probes.

Following this process of eliminating clones putatively encoding other steroid receptors, the DNA inserts in the remaining clones were analyzed by restriction mapping and subcloned into M13 vectors for di-deoxy sequence analysis. See, Chang, et al., J. Biol. Chem., 262:2826 (1987). Nucleotide sequence analysis allowed four clones to be identified as hMR-cDNA clones.

Through this stepwise process of elimination, 54 human testis clones and 6 rat prostate clones were selected and were then categorized into two groups. The first group, designated "TR2-type" cDNA comprised 30 human testis clones having sequences that overlap to form a 2.1 kb cDNA. The second group, designated "AR-type" cDNA comprised 24 human testis and 6 rat prostate clones having sequences that overlap to form a cDNA of about 2.7 kb.

EXAMPLE 2

Confirmation of the Presence on the Human X-Chromosome of an AR-type cDNA Sequence Rather than a TR2-type cDNA Sequence

The length between the putative polyadenylation signal (AATAAA) and the 5'-end in the "TR-2 type" cDNA is only 2.0 kb, which is considerably shorter than that for the cDNA of other steroid receptors. Therefore, it was suspected that the "AR-type" cDNA, rather than the "TR2-type" cDNA, encoded androgen receptor. To obtain additional information, a human X-chromosome library prepared according to Kunkel, et al., Nucleic Acids Research, 11:7961 (1983) was probed with the TR2-type cDNA and AR-type cDNA of Example 1. With TR2-type cDNA fragments, no positive clones were detected, while 3 positive clones were obtained with a 1.9 kb fragment of AR-type cDNA from a human testis (clone AR 132),

thereby confirming the presence of an AR-type cDNA sequence on the human X-chromosome. Because the X-chromosome has been implicated as the chromosome which contains an AR gene [Lyon, et al., Nature (London), 227:1217 (1970); Meyer, et al., Proc. Nat'l. Acad. Sci. (USA), 72:1469 (1975); and Amrhein, et al., Proc. Nat'l. Acad. Sci. (USA), 73:891 (1976)], this information suggested that "AR-type" cDNA, but probably not the "TR2-type" cDNA, contained the DNA sequence that could encode for androgen receptor.

Two human clones containing DNA inserts that overlapped to form a 2.7 kb cDNA were designated AR 132 and AR 5. Two rat clones containing DNA inserts that overlapped to form a 2.8 kb cDNA were designated rAR 1 and rAR 4. After restriction enzyme digestion, the DNA segments from these AR-type clones were ligated, selected and amplified using pBR322 and pGEM-3Z vectors as described in Example 3 below.

EXAMPLE 3

A. Preparation of a Human cDNA Containing AR-type cDNA from Two Different Clones and Ligation Into the Cloning Vector pGEM-3Z Plasmid

Figure 1 relates to the strategy employed in the construction of a full length hAR-cDNA clone. cDNA of clone AR 132 was digested with Eco RI to obtain a 1.9 kb fragment which was then digested with Kpn I to provide a 1 kb Eco RI-Kpn I fragment. This 1 kb fragment was ligated to a 3 kb fragment obtained by digestion of clone AR 5 with Kpn I and Pvu I. The resulting 4 kb fragment was inserted into Eco RI and Pvu I-digested pBR322 vector and used to infect *E. coli* DH5α. The transformed clones were selected by tetracycline-resistance. The plasmid with the DNA insert was digested with Cla I and Nde I to obtain a 2.6 kb fragment. The

fragment was blunt-ended with the Klenow fragment of *E. coli* DNA polymerase I and ligated to the cloning vector pGEM-3Z plasmid DNA (Promega Biotec, Madison WI.) which was previously blunt-ended by digestion with Sma I. *E. coli* DH5 α cells were transformed with the plasmid so formed (designated plasmid PhAR3600) and colonies containing the plasmid were selected by ampicillin resistance and amplified. *E. coli* DH5α cells, transformed with plasmid PhAR3600, were designated EC-hAR3600 and were deposited with the American Type Culture Collection, 12301 Parklawn Drive, Rockville, Maryland 20852 on January 25, 1989 under Accession No. 67879.

The plasmid DNA was isolated and its structure analyzed by restriction enzyme mapping and sequencing. The 2.0 kb hAR fragment obtained by NruI-BamHI digestion of a 2.6 kb hAR in pGEM3Z was then ligated to another 1.6 kb ECORI-NruI fragment of hHR to obtain the full length 3715 bp hAR. The open reading frame is about 2.8 kb which is sufficient to code for a protein with more than 900 amino acids. Near the middle of the protein is a cysteine-rich region with a 72 amino acid sequence highly homologous to regions in other steroid receptors considered to be the DNA binding domain.

As set out in detail below and illustrated in Figure 2, formation of "full length" rat AR clones by slightly varying procedures results in constructions providing RNA transcripts translatable to 79 kD and 98 kD protein products.

B. Preparation of a Rat 2.7 kb cDNA and Ligation Into the Cloning Vector pGEM-3Z Plasmid

The 2.4 kb Eco RI-Eco RI cDNA insert of clone rAR 1 was digested with Xmn I to obtain a 2.3 k b fragment. This 2.3 kb Xmn I-EcoR I fragment was ligated to a 400 bp fragment that was obtained by digestion of another cDNA clone insert (Eco RI-Eco RI insert of rAR 4) with Pst I. The ligated 2.7 kb fragment was inserted into Sma I and Pst I-digested pGEM-3Z vector and used to infect E. coli DH5α. The E. coli DH5α cells were transformed with the plasmid and colonies containing the plasmid were selected by ampicillin resistance and amplified. These cells were designated EC-rAR 2830 and were deposited with the American Type Culture Collection, 12301 Parklawn Drive, Rockville, Maryland 20852 on January 25, 1989 under Accession No. 67878. As noted in Figure 2, this construction allowed for a transcription product translated beginning with the second of two in-frame methionine-specifying codons (designated ATG_2).

C. Preparation of a Rat 2.83 kb cDNA Ligation Into the Cloning Vector pGEM-3Z Plasmid

The 2.4 kb Eco RI-Eco RI cDNA insert of rAR 1 was digested with Hind III to obtain a 1.68 kb fragment. The 1.68 kb Eco RI-Hind III fragment was ligated to a 1.15 kb DNA fragment obtained by digestion of another cDNA clone insert (rAR 6) with Hind III and Pst I. The ligated 2.83 kb fragment was inserted into Eco RI and Pst I-digested pGEM 3Z vector and used to infect E. coli DH5α. E. coli (DH5α) cells were transformed with the plasmid and colonies containing the plasmid were selected by ampicillin resistance and amplified. As noted in Figure 2, this construction allowed for a

transcription product translated beginning at the first of two in-frame methionine-specifying codons (designated ATG_1).

Figure 3 provides the nucleotide sequence of the DNA sequence of the longer "full length" rat and human AR clones and includes the deduced amino acid sequences. The first and second methionine-specifying codons are designated at amino acid positions 1 and 170 of rAR and positions 1 and 185 of hAR.

EXAMPLE 4

Transcription and Translation of the Human AR-type cDNA Plasmid in a Rabbit Reticulocyte Lysate System

pGEM-3Z vector (20 µg) containing 2.6 kb hAR DNA segment, as described in Example 3, was linearized with restriction enzyme Bam HI, phenol/chloroform extracted, and precipitated with ethanol. The linearized plasmid was transcribed in a reaction mixture containing 40 mM Tris-HCl, pH 7.5, 6 mM $MgCl_2$, 2 mM spermidine, 10 mM NaCl, 10 mM DTT, 500 µM each of ATP, GTP, CTP, and UTP, 160 units ribonuclease inhibitor, 5 µg plasmid, 30 units T7 RNA polymerase (Promega Biotec, Madison, WI) and diethylpyrocarbonate (DEPC)-treated water to a final volume of 100 µl. T7 RNA polymerase was used in the transcription of the plasmid DNA, because a T7 promotor, rather than the SP6 promotor, was found ahead of the 5'-end of the ligated AR-cDNA.

The reaction was allowed to proceed for 2 hrs. at 40°C. RQ1 DNase I (5 units) was added and the reaction continued for 15 mins. at 40°C. The reaction mixture was extracted with phenol/chloroform (1:1) and then with chloroform. RNA product was precipitated by the addition of 0.1 volume of 3 M Na-acetate and 2.5 volumes of ethanol, re-suspended in 0.5 M NaCl, and re-

precipitated with 2.5 volumes of ethanol. RNA transcribed was isolated and then translated in a rabbit reticulocyte lysate system.

Translation of RNA was carried out in a micrococcal nuclease-treated rabbit reticulocyte lysate (Promega Biotec, Madison, WI) pre-mixed kit (100 µl) in the presence of 8 µg mRNA, 40 µCi of [^{35}S] methionine (800 Ci/mmol; Amersham Co., Arlington Heights, IL) and 100 µM each of amino acid mixture without methionine. The reaction was allowed to proceed for 1 hour at 30°C. To quantitate the incorporation of radioactive methionine, 3 µl of the reaction mixture were added to 1 ml of 1 M NaOH containing 1.5% H_2O_2, 1 mM methionine, and 0.04% bovine serum albumin. The mixture was incubated for 15 mins. at 37°C to hydrolyze [^{35}S] methionine charged tRNA. The radioactive protein products were precipitated by the addition of 1 ml of 25% tricholoacetic acid and the radioactivity associated with the precipitates was determined.

By SDS-PAGE (8% acrylamide gel) analysis, performed as described in Saltzman, et al., J. Biol. Chem., 262:432 (1987), it was found that a 79 kD protein comprised more than 85% of the translated products.

EXAMPLE 5

Binding Activity of the 79 kD hAR Protein to a Synthetic Androgen

To study the steroid binding activity of the protein coded for by the cloned cDNA, the reticulocyte lysate of Example 4, containing the newly synthesized protein was incubated with 17α[^{3}H]-methyl-17β-hydroxy-estra-4,9,11-trien-3-one ([^{3}H] R1881), a potent synthetic androgen that binds AR with high affinity [Liao, et al., J. Biol. Chem., 248:6154 (1973)].

Specifically, RNA transcribed from the cloned cDNA, as described in Example 4, was translated in a rabbit reticulocyte lysate system and aliquots of the lysate were then incubated with 5 nM [^{3}H] R1881 (87 Ci/mmol) in the absence or presence of 25 nM, 50 nM, or 250 nM of non-radioactive steroid. The final incubation volume was 100 µl. The radioactive androgen binding was measured by the hydroxylapatite-filter method as described in Liao, S., et al., J. Steroid Biochem., 20:11 (1984). The result was expressed as a percentage of the radioactivity bound in the control tube (5000 dpm) without additional non-radioactive steroid and is listed in Table 1.

TABLE 1

Androgen-specific binding of hAR coded by cloned cDNA

Non-radioactive steroid added	[^{3}H] R1881-bound (% of control) 25 nM	50 nM	250 nM
R1881	13	10	1
5α-dihydrotestosterone	25	17	6
5β-dihydrotestosterone	89	89	81
17β-Estradiol	91	91	86
Progesterone	100	91	92
Dexamethasone	100	93	93
Hydrocortisone	96	90	90
Testosterone	38	28	Not tested

As shown in Table 1, the active natural androgen, 17β-hydroxy-5α-androstan-3-one(5α-dihydro-testosterone) competed well with [^{3}H] R1881 binding, but the inactive 5β-isomer did not compete well with [^{3}H] R1881 suggesting that it does not bind tightly to AR. The binding activity was steroid specific; dexamethasone, hydrocortisone, progesterone, and 17β-estradiol did not

compete well with the radioactive androgen for binding to the 79 kD protein.

Similar steroid binding specificities have also been observed for rAR coded for by cloned cDNA. Chang, C., et al., Proc. Nat'l. Acad. Sci. (USA), 85:7211-7215 (1988).

Using the hydroxylapatite filter assay method, it was observed that approximately one molecule of the ^{35}S-labelled 79 kD protein obtained from the lysate bound about one molecule of the tritiated androgen at a saturating concentration of ligand. By Scatchard plot analysis, the apparent dissociation constant was 0.31 nM, which is similar to the binding constant (0.65 nM) reported previously for AR of rat ventral prostate as reported in Schilling, et al., The Prostate, 5:581 (1984).

EXAMPLE 6

Binding Activity of the 79 kD Protein to Human Auto-antibodies

It has previously been reported [Liao, et al., Proc. Nat'l. Acad. Sci. (USA), 82:8345 (1985)] that some older men with prostate cancers have high titers of auto-immune antibodies to AR in their serum samples. The ability of human auto-antibodies to recognize the 79 kD protein made by the reticulocyte lysate system was therefore studied. The receptor protein made in the lysate system of Example 4 was incubated with [^{3}H] R1881 to allow the formation of radioactive androgen-androgen receptor (A-AR) complexes and was then mixed with serum containing auto-antibodies.

Reticulocyte lysate containing translated AR was incubated with [^{3}H] R1881, as described in Example 4, and then incubated again in either the presence of or absence of 5 µl of human male serum containing anti-

bodies to AR (anti-AR serum) for 4 hrs. at 4°C. Rabbit serum containing anti-human immuglobulins (Anti-IgG) was then added as the second antibody. After 18 hrs. of incubation at 4°C, the mixture was centrifuged and the radioactivity associated with the precipitate was estimated. Human female serum, not containing anti-AR antibody, was also used for comparison.

The results shown in Table 2 below, indicate a quantitative immunoprecipitation of the radioactive A-AR complexes in the presence of both the high titer human serum and a rabbit anti-human immunoglobulin IgG. By SDS-PAGE, it was also observed that the immunoprecipitated protein was the 79 kD protein.

TABLE 2

Anti-human immunoglobulin-dependent precipitation of hAR made by the translation of RNA transcribed from cloned cDNA

Sample incubated with [^{3}H]R1881	Anti-serum addition	Immunoprecipitable radioactivity(dpm)
AR coded by cDNA[a]	None	32
	+Anti-AR serum + Anti-IgG	8212
	+Female serum + Anti-IgG	430
	+Anti-IgG	8
Heated AR[b]	+Anti-AR serum + Anti-IgG	42
BMW-lysate[c]	+Anti-AR serum + Anti-IgG	204

[a] 8500 dpm of the radioactive AR complexes made were used.

[b] Reticulocyte lysate containing AR was heated at 50°C for 20 mins. to inactivate receptor and release the radioactive androgen bound before the addition of antiserum.

[c] Brome Mosaic Virus RNA was used in the reticulocyte lysate translation system instead of RNA transcribed from cloned cDNA.

EXAMPLE 7

Characterization of "TR2-type" cDNA

Of the more than 40 TR2-type human cDNA clones obtained, including the 30 described in Example 1, the clone designated TR2-5 was found to be 2029 base pairs in length as indicated in Figure 4. The open reading frame between the first ATG and terminator TAA can encode 483 amino acids with a calculated molecular weight of 52 kD. The putative DNA binding region is underscored. The putative initiator ATG matched closely with Kozak's concensus sequence for active start codons. [See, Kozak, M., Nature, 308:241 (1984).] Two triplets upstream of this ATG codon is an in-frame terminator (TAA) further supporting initiator function for the ATG.

Eleven of the 30 TR2-type clones of Example 1, as represented by the clone designated TR2-7, contain an internal 429 bp insertion between nucleotide sequence 669 and 670 (designated by an asterisk in Figure 4). This internal insertion introduces a termination codon TAG (underscored in the insert sequence footnote) which reduces the open reading frame to 184 amino acids with a calculated molecular weight of 20 kD. It is likely that the insertion in these 11 TR2 clones (or deletion in the 19 other TR2 clones) represents either the existence of two types of mRNA in the human testis or an artifact of cDNA construction. In the 3'-nontranslated region, a eukarotic polyadenylation signal AATAAA is present between the nucleotide sequence 2000 and 2007 of the TR2-5 clone.

TR2-9 receptor cDNA was isolated from a human prostate cDNA library has 1785 bp (Figure 5). The open reading frame from the first ATG to TAA encoded 467 amino acids with a calculated molecular weight of 50,849 daltons.

TR2-11 receptor cDNA has 2221 bp, with a shorter 5'-untranslated region (Figure 6). The open reading frame encoded a polypeptide of 603 amino acids with a calculated molecular weight of 67,223 daltons. The predicted initiator ATG of these two cDNA sequences matches well with Kozak's consensus sequence for an active start codon (Kozak, M., Nature, 308:241-246 (1984)) and there is an in-frame stop codon TAG upstream of the initiation ATG in each cDNA sequence. In the 3'-un-translated region, a eukaryotic polyadenylation signal AATAAA is present between nucleotide numbers 1710-1715 for the TR2-9 receptor and between 2180-2185 bp for the TR2-11 receptor.

Other variants of TR-2 with open reading frames at the putative ligand-binding domains may code for receptors for new hormones or cellular effectors. It is anticipated that the knowledge of TR2-cDNA sequences will be utilized in isolation and structural analysis of other cellular receptors, their genes, and ligands (endogenous or therapeutic agents) that can regulate cellular growth and functions in both normal and diseased organs.

There is a conservation of the DNA-binding domain for TR2 receptors and for other members of the steroid hormone receptor family. The putative DNA-binding domain of TR2 receptor shares 50-60% homology with that of other steroid receptors and TR3 receptor (Chang, C., Kokontis, J., and Liao, S., Science, 240:324-326 (1988); Chang, C., Kokontis, J., Chang, C.T., and Liao, S., Nucleic Acid Res., 22:9603 (1987); Green, S., Walter, P., Kumar, V., Krust, A., Bornert, J.M., Argos, P., and Chambon, P., Nature, 320:134-139 (1986); Arriza, J.L., Weinberger, C., Cerelli, G., Glaser, T.M., Handelin, B.L., Housman, D.E., and Evans, R.M., Science, 237:268-275 (1987)). TR3 receptor is another member of the steroid receptor family, which may

be a human homologue of the mouse NUR/77 gene product (Chang, C., Lau, L., Liao, S., and Kokontis, J., in the Steroid/Thyroid Hormone Receptor Family and Gene Regulation, Birkhauser Verlag, Basel, Boston, Berlin, pp. 183-193 (1988); Hazel, T.G., Nathans, D., and Lau, L.F., Proc. Nat'l. Acad. Sci. USA, 85:8444-8448 (1988)). The 26 amino acids in the DNA-binding domain of TR2 receptor are identical to those in the DNA-binding domain of all other known steroid receptors. The positions of conserved amino acid residues have been proposed to be involved in the formation of DNA-binding domain "Zinc fingers" (Weinberger, C., Hollenberg, S.M., Rossenfeld, M.G., and Evans, R.M., Nature, 318:670-672 (1985)).

Figure 7 depicts an amino acid sequence alignment of the cysteine-rich DNA binding domain of human androgen receptor, glucocorticoid receptor, mineralocorticoid receptor, progesterone receptor, estrogen receptor, human TR2 protein, rat AR, chick vitamin D receptor (c-VDR), and the v-erb A oncogene product of avian erythroblastosis virus. The numbers in the left margin represent the positions of amino acid residues in the individual receptors. Common residues are boxed with solid lines. The residues in dotted boxed represent those not in common with those in the solid boxes. V-erb A has two more amino acids at the starred position.

In this region, the human and rat cDNAs for AR have identical amino acid sequences, although for some amino acids different codons are employed. Also in this region, the homology between human AR or rat AR and other receptors is as follows: glucocorticoid receptor (GR), 76.4%; mineralo-corticoid receptors (MR), 76.4%; progesterone receptors (PR), 79.2%; estrogen receptors (ER), 55.6%; TR2, 45.8%; chick vitamin D receptor (c-VDR), 40,3%; and the v-erb A oncogene product of avian

erythroblastosis virus, 40.3%. In the putative region for steroid binding, which has about 200 amino acids near the -COOH terminal of steroid receptors, the homology between human AR or rat AR and hGR, hMR, or hPR is about 45-55%, whereas the homology between human AR and rat AR and hER is less than 20%. Thus, human and rat AR appear to be more closely related to GR, MR, and PR, than to v-erb A or to receptors for estrogen, vitamin D, and thyroid hormones.

The DNA binding domain of TR2 (amino acids 111 to 183) has a high homology with the steroid receptor super-family as follows: retinoic acid receptor (RAR), [Giguere, et al., Nature, 330:624 (1987)], 65%; thyroid receptor (T_3R) [Sap, et al., Nature, 324:635 (1987)], 59%; mineralocorticoid receptor (MR), [Arriza, et al., Science, 235:268 (1987)], 54%; vitamin D_3 receptor (VD_3R) [McDonnell, et al., Science, 235:1214 (1987)], 53%; hERR1 and hEER2, [Giguere, V., et al., Nature, 331:91 (1988)], 51% estrogen receptor (ER), [Hollenberg, et al., Nature, 318:635 (1985)], 51%; glucocorticoid receptor (GR) [Hollenberg, et al., Nature, 318:635 (1985)], 50%; androgen receptor (AR), 50%; progesterone receptor (PR), 49%; [Loosfelt, et al., Proc. Nat'l. Acad. Sci.,(USA), 83:9045 (1986)]. As noted in Figure 7, the positions of 20 amino acids (9 Cys, 3 Arg, 2 Gly, 2 Phe, 1 Lys, 1 Met, 1 Asp, 1 His) in the putative DNA binding domain are identical among all isolated thyroid steroid receptor genes. It has been proposed that this highly conserved region may be involved in the formation of a DNA binding finger. See, Weinberger, et al., Nature, 318:670 (1985). Like the other steroid receptors, TR2 does not have the two extra amino acids (Lys-Asn) found only in the thyroid receptors' DNA binding domain. See, Sap, et al., Nature, 324:635 (1987).

EXAMPLE 8

In Vitro Transcription and Translation of TR2 cDNA

The Eco RI-Eco RI DNA inserts from clones TR2-5 and TR2-7 were isolated and ligated to an EcoR1 digested pGEM-3Z vector for in vitro transcription essentialy as described in Example 3. E. coli DH5α cells, transformed with these plasmids were designated EC TR2-5 and EC TR2-7 and were deposited January 25, 1989 with the American Type Culture Collection, 12301 Parklawn Drive, Rockville, Maryland 20852 under Accession Nos. 67877 and 67876.

Transcribed RNA was then translated in a rabbit reticulocyte lysate system. By SDS-polyacrylamide gel electrophoresis (PAGE), it was found that the major translated product of TR2-7, which has an internal 429 bp, insertion, was a 20 kD protein. The major translated product of TR2-5 was a 52 kD protein.

TR2-11 receptor cDNA was isolated and ligated to EcoRI-digested pGEM-3Z vector for in vitro transcription, essentially as described in Example 3. E. coli DH5α cells, transformed with this plasmid, were designated EC TR2-11 and deposited on November 14, 1989; with the A.T.C.C. under accession No. 68173. Transcribed RNA was translated in a rabbit reticulocyte lysate system. SDS polyacrylamide gel analysis showed a major band around 67 kd, consistent with the calculated molecular weight of 67,223 daltons.

To further characterize these translated proteins, the translation lysate was passed over a DNA cellulose column. The bound product was then eluted, concentrated and applied to SDS-PAGE. The results indicated that the translated proteins were indeed DNA-binding proteins.

EXAMPLE 9

Binding Activity of TR2-5, TR2-7 and TR2-11 cDNA Expression Product

To study the steroid binding activity of the translation products of the TR2-5, TR2-7, and TR2-11 clones, the products were incubated with all major classes of steroids, including androgens, progesterone, glucocorticoid and estrogen but no significant binding with the above steroids was observed. This does not necessarily rule out a steroid binding function for these proteins. Possibly the TR2-5, TR2-7, and TR2-11 expression products' steroid binding activity may involve some post-translation modifications missing in the rabbit reticulocyte lysate system. Alternatively, the TR2-5, TR2-7, and TR2-11 translated proteins may be steroidal independent or may bind to an unidentified ligand present in the human testis or rat ventral prostate, or, alternatively, may be dependent upon an unknown steroidal or non-steroidal hormone.

The size of TR2 mRNA was determined by Northern blot analysis with TR2-5 cDNA insert as a probe. One 2.5 kb band was detected which should include enough sequence information to code for a 52 kD protein. The TR2 mRNA tissue distribution was also analyzed by dot hybridization. The hybridization was visualized by densitometric scanning of the autoradio-graphs, individual dots were cut and radioactivity measured by liquid scintillation counting. Chang, et al., J. Biol. Chem., 262:2826 (1987). The results showed that TR2 mRNA was most abundant in the rat ventral prostate with the relative amounts in other tissues being: prostate 100%, seminal vesicle 92%; testis, 42%; submaxillary gland, 18%; liver, 13%; kidney, <1%; and uterus, <1%.

EXAMPLE 10

Schematic Comparison of the Four Variants of Human TR2 receptors:

A schematic comparison of four TR2 receptors (TR2-5; TR2-7; TR2-9; and TR2-11) is shown in Figure 8. TR2-7 receptor contains an internal extra 429 base point segment between nucleotide number 670 and 671 base point, which generates a termination codon and shortens the open reading frame to 184 amino acids. Chang, C., Kokontis, J., B.B.R.C., 155:971-977 (1988).

The sequences of TR2-5, TR2-9, and TR2-11 receptors are identical from amino acid number 1 to 464. However, the C-terminal hormone-binding domains of these three TR2 receptors are different. Chang, C., Kokontis, J., B.B.R.C., 155:971-977 (1988). TR2-9 receptor has 16 fewer amino acids and 3 different amino acids as compared with TR2-5 receptor, due to a 244 bp insertion between nucleotide number 1518 and 1763 of TR2-5 receptor. Evans, R.M., Science, 240:889-894 (1988). TR2-11 receptor has more and quite different amino acids in the hormone-binding domain.

The variant forms of TR2 receptors, like multiple forms of thyroid hormone receptors, (Evans, R.M., Science, 240:889-894 (1988)), may be very significant in terms of biological function. However, there are differences with respect to tissue specificity and with respect to the degree of homology in the putative DNA-binding domain. Variant thyroid hormone receptors were found in different tissues, indicating tissue specificity of the receptors. In contrast, although TR2-11 receptor cDNA was isolated from human prostate cDNA library, all other TR2 receptor cDNAs (TR2-5, TR2-7, and TR2-9) were isolated from a human testis cDNA library, indicating co-expression in at least one human tissue. The incomplete homology in the

DNA-binding domain of thyroid receptors may contribute to the differential target gene specificity. In contrast, the putative DNA-binding domain of TR2 receptors are identical, suggesting that they may act on the same target gene(s). Variant TR2 receptors may be the products of different genes. Alternatively, RNA splicing can generate messages encoding TR2 receptors with multiple hormone-binding domains. If this is the case, regulation at the RNA splicing level may be important during the transition of hormone-dependent organs/tumors to hormone-independent organs/tumors. Also, if TR2 receptors with different hormone-binding domains are able to bind to different natural ligands, or to the same ligand with a differential affinity, the co-expression of variant receptors may provide competition for ligands among receptors, and the activation level of the target genes could be regulated by adjusting the expression ratio of different variant receptors. This expression ratio could vary with tissue-specificity or developmental stage-specificity. Given that in rat, TR2 receptor mRNA was most abundant in the androgen-sensitive ventral prostate (Chang, C., Kokontis, J., B.B.R.C., 155:971-977 (1988)), it is of interest to examine the expression ratio of variant TR2 receptors in normal, neoplastic, or hyperplastic prostate tissue and study their possible roles in prostate growth and development. It is anticipated that a determination of the genomic structure of TR2 receptor genes and the natural TR2 receptor ligand may lead to elucidation of the mechanism by which variant receptors are generated and elucidation ofthe cellular function of this new member of the steroid hormone receptor superfamily.

EXAMPLE 11

Analysis of Androgen Regulation of AR and TR2 mRNA Levels in the Rat Ventral Prostate

Because rat ventral prostate is an androgen-sensitive organ and contains the greatest amount of AR and TR2 mRNA, the effect of androgen depletion and replacement on the mRNA levels was studied by RNA dot hybridization and Northern blot analysis. Total RNA was extracted from the ventral prostate of normal rats, rats castrated and rats previously castrated and treated with 5α-dihydrotestosterone (17β-hydroxy-5α-androstand-3-one). AR mRNA levels per unit of DNA increased 200 to 300% of the level for normal rats within 2 days after castration. Administration of 5α-dihydrotestosterone (5 mg/rat/day) into castrated rats reduced the AR mRNA level to that of normal rats. TR2 mRNA levels, per unit of DNA, were increased to 170% of the normal rat within 2 days after castration. Injection of 5α-dihydrostestosterone (5 mg/rat/day) into castrated rats reduced the TR2 mRNA to the levels of normal rats. Interestingly, the total prostate RNA levels, at the same period of time, were decreased to 40% of the normal level. The effects of androgen on the levels of prostatic TR2 mRNA were further confirmed by flutamide injection experiments. Flutamide, an anti-androgen which antagonizes the effects of 5α-dihydrotestosterone on the ventral prostate weights in castrated rats [Neri, et al., Invest. Urol., 10:123 (1972)], was injected into normal rats for from 2 to 6 days. TR2 mRNA levels were then measured by dot hybridization as described above. The results show that flutamide injection, like castration, increased TR2 mRNA levels. The change in the AR or TR2 protein levels could be due to a change in mRNA stability and utilization or a change in the regulation of gene transcription. The activation or inactivation by

androgen of specific genes to different degrees in the same organ may suggest that androgen is involved in the structuring of the pattern of gene expression in the target cell. Also, if androgen-mediated gene repression mechanisms are related to growth of the prostate, then a further study of the mechanism and structure of genes, repressed AR and TR2 mRNA may provide a better understanding of androgen action in the normal and abnormal prostate and other hormone sensitive organs.

Also, defects in the structures of AR and androgen sensitive genes and/or losses of the control of the production and function of these gene products can be the causes of the abnormal growth of androgen sensitive or insensitive tumors like prostate cancers. These lines of research may, therefore, be helpful in designing new diagnostic methods and treatments for patients.

EXAMPLE 12

Expression of Cloned AR-Genes and Androgen Sensitive Genes in Eukaryotic and Prokaryotic Cells

The ability of cloned genes to function when introduced into mammalian, yeast, and bacterial cells has proved to be very valuable in understanding the function and regulatory mechanism of genes. Recombinant techniques can provide, in large quantities, gene expression products (proteins) which are not readily obtainable from natural sources. While bacterial systems are very useful in large scale production of those proteins which do not require substantial post-translational modification for optimal biological activity, eukaryotic systems are particularly advantageous because of their ability to correctly modify the expressed proteins to their functional forms.

Using well known techniques, AR-cDNA and TR2-cDNA may readily be used for large scale production of gene products. For this purpose, the most efficient transcription units can be constructed using viral, as well as non-viral, vectors with regulatory signals that can function in a variety of host cells. SV40, pSV2, adenoviruses, and bovine papilloma virus DNA have been used successfully for introduction of many eukaryotic genes into eukaryotic cells and permit their expression in a controlled genetic environment. These and similar systems are expected to be appropriate for the expression of AR- and TR2-genes. To assist gene transfer, the two most widely used methods, the "calcium phosphate precipitation" and the "DEAE-dextran technique" can be used. Genes can be introduced into cells either transiently, where they continue to express for up to 3 days, or, more permanently to form stably transformed cell-lines. The expressed proteins can be detected by androgen binding or antibody assays.

The expression of cloned AR-genes was achieved as follows in a eukaryotic system. NIH 3T3 cells, contact-inhibited cells established from NIH Swiss mouse embryo, were co-transfected with hAR cDNA inserted into pBPVMTH vectors as described by Gorman, "DNA Cloning", 2:143-190 D. M. Glover, ed.; (Oxford, Washington, D.C. 1985). Transfected cells were cloned and grown in multiple-well cell culture plates. About 100 individual cell lines were isolated. Of these, 6 demonstrated [^{3}H] R1881-binding activity at least 4-fold the activity of cells transfected with pSV2 vector alone, i.e., without the hAR cDNA sequence.

To express AR cDNA in prokaryotic systems, hAR and rAR cDNAs were inserted into a number of expression vectors including pUR, λGT11, pKK223-3, pKK233-2, pLEX, pATH1, pATH2, pATH10, and pATH11. Vectors with AR cDNA inserts were used to infect *E. coli* strains (JM109,

DH5α, Y1089, JM105, and RR1). According to polyacrylamide gel electrophoresis analysis, the infected bacteria can synthesize AR fragments coded for by the AR cDNA inserts. Some of these AR polypeptides are degraded in culture. Amino terminal, DNA-binding, and androgen binding domains were used, as described in Example 13, to construct fusion proteins representing these domains.

EXAMPLE 13

Production of Polyclonal and Monoclonal Antibodies to AR

The isolation of AR in significant amounts from androgen sensitive organs has been exceedingly difficult. Therefore, the high-level expression of hAR or rAR cDNAs, as shown in Example 12, is expected to be an ideal way for the large scale production of AR. In addition, oligopeptides, with sequences identical to the deduced amino acid sequences of portions of AR molecules, can be chemically synthesized inexpensively in large quantities. Both AR produced by expression vectors in eukaryotic or prokaryotic cells and AR oligopeptides chemically synthesized were used as antigens for the production of monoclonal antibodies as described in greater detail below.

Generally, several chemically synthesized oligopeptides, representing sequences unique to AR, (i.e., PYGDMRLETARDHVLP; CPYGDMRLETARDHVLP; and SIRRNLVYSCRGSKDCIINK) were bound to BSA or KLH carrier proteins and were used to immunize mice. Spleen cells from these mice were fused to myeloma cells to produce hybrid antibody producing cells. Analysis by ELISA (enzyme-linked immunoassay) of the supernatants of 4 hybrid cultures appeared to indicate the presence of immunoglobulin that interacts with AR of rat ventral

prostate. It is anticipated that these cells which produce monoclonal antibodies can be injected intraperitoneally into BALB/c mice previously treated with pristane. Ascites fluids can then be harvested and antibodies precipitated with ammonium sulfate.

Expression of Androgen Receptor Fusion Protein in E. coli

Three different parts of the AR gene (encompassing the N-terminal domain, the DNA-binding domain and the androgen-binding domain) were fused, in frame, to the N-terminal half of the trpE gene (trpE promoter-the first 969 bp of trpE coding region-multiple cloning region of pUC12) by using the pATH expression vectors as shown in Figures 9, 10, and 11; respectively. Dieckmann, *et al*., *J. Biol. Chem.*, *260*:1513 (1985).

These constructions resulted in the fusion of approximately 25 kDa of AR, including a portion of the N-terminal domain; 29 kDa of AR, including a major portion of the DNA-binding domain; and 12 kDa of AR, including a portion of the androgen-binding domain; to the 33 kDa trpE protein. Because the trpE protein is insoluble, partially purified induced fusion proteins were obtained simply by lysing the *E. coli* and precipitating the insoluble fusion proteins. After electrophoresis on SDS-polyacrylamide gels, the induced fusion proteins, i.e., those proteins not present in the control pATH vector (no AR gene insert), were sliced from the gels and then used for immunization.

Fusion proteins, other than the three specifically exemplified, can also be constructed using these means.

Production and Purification of Anti-AR Antibodies

Rabbits, rats, and mice were immunized with either SDS-polyacrylamide gel slices containing denatured fusion proteins or electro-eluted, SDS-free, fusion protein, as well as fusion proteins obtained by other protein purification methods. The presence of antibodies to the fusion proteins in the antisera was assayed by ELISA. Positive serum having a higher titer was further assayed by the double antibody precipitation method using rat ventral prostate cytosol [^{3}H]AR as antigen. The results showed that 1 µl of crude serum precipitated 10 to 20 fmole [^{3}H]AR. Anti-AR crude serum was then affinity-purified by differential suspension of immune serum containing TrpE protein(s) (both those TrpE proteins having and those TrpE proteins not having inserted AR sequences) expressed by pATH vectors. The bound antibodies can be removed from the suspension because TrpE protein is insoluble. Antibodies specific against only the trpE protein were removed; antibodies specific for AR were isolated and again confirmed by both ELISA and double antibody precipitation.

Production of Monoclonal Anti-Androgen Receptor Antibodies

The immunized rats were judged ready to be sacrificed for a fusion when their serum tested positive anti-AR antibodies by ELISA. Spleens were removed and grinded to release the cells into DMEM (Dulbeco's Modified Engle's Medium) medium. Through a series of centrifugations using DMEM + DMEM with Ficoll Hypaque, the spleen cells were isolated. The SP2/0 myeloma cells were grown, split and diluted in 50 ml of DMEM with 20% FCS, 1% MOPS, and 1X L-Gln for two days before ready for the fusion. SP2/0 cells (5 x 10^6) and 5 x 10^7 spleen cells were used in the fusion. After incubating overnight, the fused cells were collected, suspended in

DMEM with 1X H-T, 1X Methotrexate, 20% FCS, and 1X PBS and distributed in 96-well plates. Plates were supplemented after 6 days with DMEM and 20% FCS. Hybridomas were identified and assayed, using the ELISA assay of Engrall, et al., Bio. Chem. et Biophys. ACTA, 251:427-439 (1971). In this assay, plates were coated with either the AR fusion proteins or the TrpE protein as antigen and read on an ELISA reader.

Only those hybridomas that caused a positive reaction with the AR fusion protein were "limit diluted" to a concentration of 10 cells/ml and were then distributed among half of a 96-well plate. The remaining cells from the original well were transferred to a 24-well plate. Each of these plates had a thymocyte feeder layer. The tymocyte feeder layer was made up of thymus cells isolated from an un-injected rat, purified through centrifugation, irradiated with 1200 to 1400 RADS, and diluted to 1 x 10^7 cells/ml of DMEM with 20% FCS.

Positives from these thymocyte 96-well plates were again tested by ELISA. Only those which again tested positive with the AR fusion protein were grown up for monoclonal antibody purification. Three of the wells produced monoclonal antibody against AR. Both ELISA and double antibody assays were positive. The monoclonal antibodies were designated AN1-6, AN1-7, and AN1-15 and the three cell lines were designated HAN1-6, HAN1-7, and HAN1-15; Accession Nos. 10,000; 9,999; and 10,001; respectively, deposited on January 25, 1989 with the American Type Culture Collection, 12301 Parklawn Drive, Rockville, Maryland 20852.

Specificity of Anti-AR Antibodies

Sucrose gradient centrifugation was used to characterize the specificity of the three monoclonal anti-AR antibodies and their ability to react with non-denatured [^{3}H]AR.

Cytosol was prepared from the ventral prostates of castrated rates as follows. Rats were castrated by the scrotal route while under anesthesia. They were killed 18 hrs. laters by cervical dislocation and their ventral prostates were removed, minced with scissors, washed in Buffer A (50 mM sodium phosphate, pH 7.5, 1 mM EDTA, 2 mM DTT, 10 mM sodium molybdate, 10% (v/v) glycerol and 10 mM sodium floride) and homogenized in 2x the tissue volume of Buffer A + 0.1 mM bacitracin, 1 mM PMSF, and aprotinin (1TIU/ml). The homogenate was centrifuged at 5,000 x g for 10 mins., adjusted to 10 nM ^{3}H-androgen, spun at 225,000 x g for 45 mins. and treated with dextran-coated charcoal. One hundred µl of the cytosol solution, containing 3H-A-AR complexes, was incubated for 6 hrs. with 100 µl of the purified anti-androgen receptor monoclonal antibody, AN1-6, (20x as concentrated as the tissue culture media). Sucrose gradient centrifugation was performed by centrifugation at 257,000 x g for 16 hrs. at 4°C on a 3.8 ml, linear 5-20% (w/v) sucrose gradient containing 20 mM Tris-HCl, pH 7.5, 1 mM EDTA, 1 mM DTT, 10% (v/v) glycerol, and 0.4 M KCl. Gradients were fractionated and numbered from the bottom and 0.2 ml per fraction collected. The results obtained indicated that all three of the monoclonal antibodies, AN1-6, AN1-7, and AN1-15, recognized and effectively bound the radioactively labeled androgen receptor ([^{3}H] AR).

The [^{3}H]AR and other steroid receptor complexes had a sedimendation coefficient of about 4-5S in the sucrose gradient media containing 0.4M KCl. Anti-AR antibodies do not alter the sedimentation coefficient of 4-5S for [^{3}H]glucocorticoid receptors complexes of rat liver, estrogen receptor complexes of MCF-7 cells, and progesterone receptor complexes of T47D cells, but do shift the sedimentation coefficient of [^{3}H]A-AR complexes of rat ventral prostate from 4S to 9-

12S or to heavier units. By SDS-polyacrylamide gel electrophoresis analysis it was also found that all major <u>in vitro</u> transcription/translation products of human and rat AR cDNAs were immunoprecipitatable by the anti-AR antibodies.

EXAMPLE 14

Use of AR cDNA and TR2 cDNA as Probes in the Study of Abnormality in Human and Animal Organs and Cancer Cells

Patients with metastatic prostatic cancer initially often respond favorably to androgen withdrawal therapy (castration or antiandrogen treatments). Most patients, however, eventually relapse to an androgen-state for which no chemotherapy, which would significantly increase the survival rate, is available. Regardless of the origin of androgen -independent or -insensitive cancer cells, it is important to understand whether the androgen insensitivity or abnormality in the diseased cells are due to qualitative or quantitative changes in (a) the AR or TR2 genes, (b) regulation of their transcription, or translation, or (c) other cellular factors. AR cDNA, TR2 cDNA, or their partial segments can be used as specific probes in these studies.

For the analysis of AR or TR2 genes, high molecular weight genomic DNA isolated from target organs, tumors, and cultured cells can be used in identifying and characterizing AR genes. Different restriction endonucleases can be used to cleave DNA. The fragments can be analyzed by Southern analysis (agarose electrophoresis, transfer to nitrocellulose and hybridization with AR cDNA probes). After identification, selected fragments can be cloned and sequenced. It is also possible to use appropriate oligonucleotide fragments of AR or TR2 cDNA as primers

to amplify genomic DNA isolated from normal and abnormal organs or cells by specific DNA polymerases. The amplified genomic DNA can then be analyzed to identify sequence abnormality using the polymerase chain reaction (PCR) assay. Saiki, et al., Science, 230, 1350 (1985). See also, Mullis, K.B., U.S. Patent No. 4,683,202; July 28, 1987; and Mullis, K.B., U.S. Patent No. 4,683,195; July 28, 1987. For the analysis of mRNA for ARs or related proteins, dot hybridization and Northern hybridization analysis could be used to characterize mRNA and AR or receptor-like molecules quantitatively and qualitatively. From these studies valuable information about the number of different forms of AR genes and their expression in androgen insensitive and sensitive tumor cells can be obtained.

DNAs and RNAs obtained from androgen sensitive and insensitive tumors and from cell lines from rats and humans with testicular feminization syndromes have been analyzed by the above methods. Preliminary studies indicated that abnormality in androgen responses may be due to sequence deletion/mutation in genes for ARs.

EXAMPLE 15

Development of Transgenic Animals

Transgenic techniques have been employed for expression of exogenous DNA. It may therefore be possible to confer androgen sensitivity to animals with androgen receptor defects. For example, androgen insensitive animals, such as testicular feminized mice or rats, are known to have defective AR genes or defective AR itself. If DNA containing a normal AR gene is injected into fertilized mouse embryos, the transgenic mice may carry and express the gene and produce a functional AR necessary for androgen responses. For micro-injection, it is necessary to use AR genes

containing DNA that can be expressed in the insensitive animals.

A number of genomic receptor clones from human X-chromosome libraries and rat genomic DNA libraries have been obtained and analyzed for their structures. Clones containing AR sequences will be characterized by endonuclease mapping, by Southern hybridization and by S1-nuclease mapping. The 5' and 3' untranslated regions thus identified will aid in determining the minimal size of the DNA that would be required for tissue specific expression of the AR coding region.

Partial sequence analysis of the 5' and 3' regions would locate the minimal region that represents the promoter and the polyadenylation region. Approximately 2 to 5 kb of upstream un-translated region and 0.5 to 1 kb of sequences downstream from the poly(A) site may be fused to the cDNA clone (minimal-gene) and injected into embryos of mice. Transgenic mice would be identified by analysis of their tail DNA using mini-gene specific probe(s).

Normally only some of the transgenic mouse lines can express their transgenes. Transgenes may be inactive because of the presence of inhibitory sequences, integration of the exogenous gene into a transcriptionally inactive chromosomal location, or the juxtaposition of the transgene and an endogenous enhancer. In addition, androgen insensitivity may be due to various other factors and not due to abnormality in the AR gene or its expression.

The foregoing illustrative examples relate to the isolation of human and rat cDNAs encoding DNA binding proteins including androgen receptor and TR-2 and more particularly describe the transcription of the corresponding cDNAs and translation of the corresponding mRNAs in cell-free systems. While the present invention has been described in terms of specific methods and

compositions, it is understood that variations and modifications will occur to those skilled in the art upon consideration of the present invention.

Accordingly it is intended in the appended claims to cover all such equivalent variations which come within the scope of the invention as claimed.

WHAT IS CLAIMED IS

1. A purified and isolated DNA sequence encoding androgen receptor polypeptide.

2. The DNA sequence according to claim 1 encoding human androgen receptor polypeptide.

3. The DNA sequence according to claim 1 encoding rat androgen receptor polypeptide.

4. The DNA sequence according to claim 1 and as set forth in Figure 3.

5. The DNA sequence according to claim 1 which is a cDNA sequence.

6. The DNA sequence according to claim 1 which is a genomic DNA sequence.

7. The DNA sequence according to claim 1 which is a partially synthetic DNA sequence.

8. A purified and isolated DNA sequence encoding TR2 polypeptide.

9. The DNA sequence according to claim 8 which is a cDNA sequence.

10. The DNA sequence according to claim 8 which is a genomic DNA sequence.

11. The DNA sequence according to claim 8 which is a partially synthetic DNA sequence.

12. The DNA sequence according to claim 8 encoding TR2-5 and as set forth in Figure 4.

13. The DNA sequence according to claim 8 encoding TR2-7 and as set forth in Figure 4.

14. The DNA sequence according to claim 8 encoding TR2-9 and as set forth in Figure 5.

15. The DNA sequence according to claim 8 encoding TR2-11 and as set forth in Figure 6.

16. A procaryotic or eucaryotic host cell transformed or transfected with a DNA sequence according to claim 1 or 8.

17. The procaryotic transformed host cell according to claim 16 which is E. coli DH5α cells designated as, and corresponding to A.T.C.C. deposit Nos.: EC-hAR 3600, A.T.C.C. No. 67879; EC-rAR 2830, A.T.C.C. No. 67878; EC TR2-5, A.T.C.C. 67877; EC TR2-7, A.T.C.C. No. 67876; and EC TR2-11, A.T.C.C. No. 68173.

18. A viral or circular DNA plasmid comprising a DNA sequence according to claim 1 or 8.

19. A viral or circular DNA plasmid according to claim 18 further comprising an expression control DNA sequence operatively associated with said DNA sequence.

20. A method for the production of androgen receptor polypeptide comprising:

growing, in culture, a host cell transformed or transfected with a DNA sequence according to claim 1; and

isolating from said host cell or culture the polypeptide product of the expression of said DNA sequence.

21. A method for the production of androgen receptor polypeptide comprising:

disposing a DNA sequence according to claim 1 in a cell free transcription and translation system; and

isolating from said system the polypeptide product of the expression of said DNA sequence.

22. A method for the production of TR2 polypeptide comprising:

growing, in culture, a host cell transformed or transfected with a DNA sequence according to claim 8; and

isolating from said host cell or culture the polypeptide product of the expression of said DNA sequence.

23. A method for the production of TR2 polypeptide comprising:

disposing a DNA sequence according to claim 8 in a cell free transcription and translation system; and

isolating from said system the polypeptide product of the expression of said DNA sequence.

24. The polypeptide product of the in vitro or in vivo expression of a DNA sequence according to claim 1.

25. An amino acid sequence as set out in Figure 3.

26. The polypeptide product of claim 24 characterized by molecular weights of 98 kD and 79 kD by SDS-PAGE and the ability to bind an androgen.

27. The polypeptide product of the in vitro or in vivo expression of a DNA sequence according to claim 8.

28. TR2 polypeptides.

29. An amino acid sequence as set out in Figure 4 and comprising TR2-5.

30. An amino acid sequence as set out in Figure 4 and comprising TR2-7.

31. An amino acid sequence as set out in Figure 5.

32. An amino acid sequence as set out in Figure 6.

33. A synthetic peptide duplicative of a sequence of amino acids present in androgen receptor or TR2 proteins in a region of the proteins not involved with DNA binding functions and sharing at least one antigenic epitope with androgen receptor or TR2 proteins.

34. An antibody specifically immunoreactive with at least one epitope of androgen receptor polypeptide or TR2 polypeptide other than an epitope within the DNA binding functional region thereof.

35. The monoclonal antibody according to claim 34.

36. The monoclonal antibody according to claim 34 and produced by hybridoma cell line Nos. HB 10,000; HB 9,999; and HB 10,001.

37. The polyclonal antibody according to claim 34.

38. A method for quantitative detection of androgen receptor based on the immunological reaction of androgen receptor with an antibody according to claim 34.

39. A method for quantitative detection of TR2 receptor based on the immunological reaction of TR2 receptor with an antibody according to claim 34.

40. A method for the quantitative detection of androgen receptor encoding DNA or RNA based on hybridization of said nucleic acids with a DNA sequence according to claim 1.

41. A method for the quantitative detection of TR2 receptor encoding DNA or RNA based on hybridization of said nucleic acids with a DNA sequence according to claim 8.

42. A method for the quantitative and qualitative detection of AR or TR2 specific gene sequence or sequences present in a sample comprising the steps of:

a) treating said sample with one oligonucleotide primer for each strand for said specific sequence, under hybridizing conditions such that for each strand of each sequence to which an oligonucleotide primer is hybridized an extension product of each primer is synthesized which is complementary to each nucleic acid strand, wherein said primer or primers are selected so as to be sufficiently complementary to each strand of each specific sequence to hybridize therewith such that

the extension product synthesized from one primer, when it is separated from its complement, can serve as a template for synthesis of the extension product of the other primer;

b) treating the sample under denaturing conditions to separate the primer extension products from their templates if the sequence or sequences to be detected are present;

c) treating the sample with oligonucleotide primers such that a primer extension product is synthesized using each of the single strands produced in step (b) as a template, resulting in amplification of the specific nucleic acid sequence or sequences if present;

d) adding to the product of step (c) a labeled oligonucleotide probe for each sequence being detected capable of hybridizing to said sequence or a mutation thereof; and

e) determining whether said hybridization has occurred.

FIGURE 1

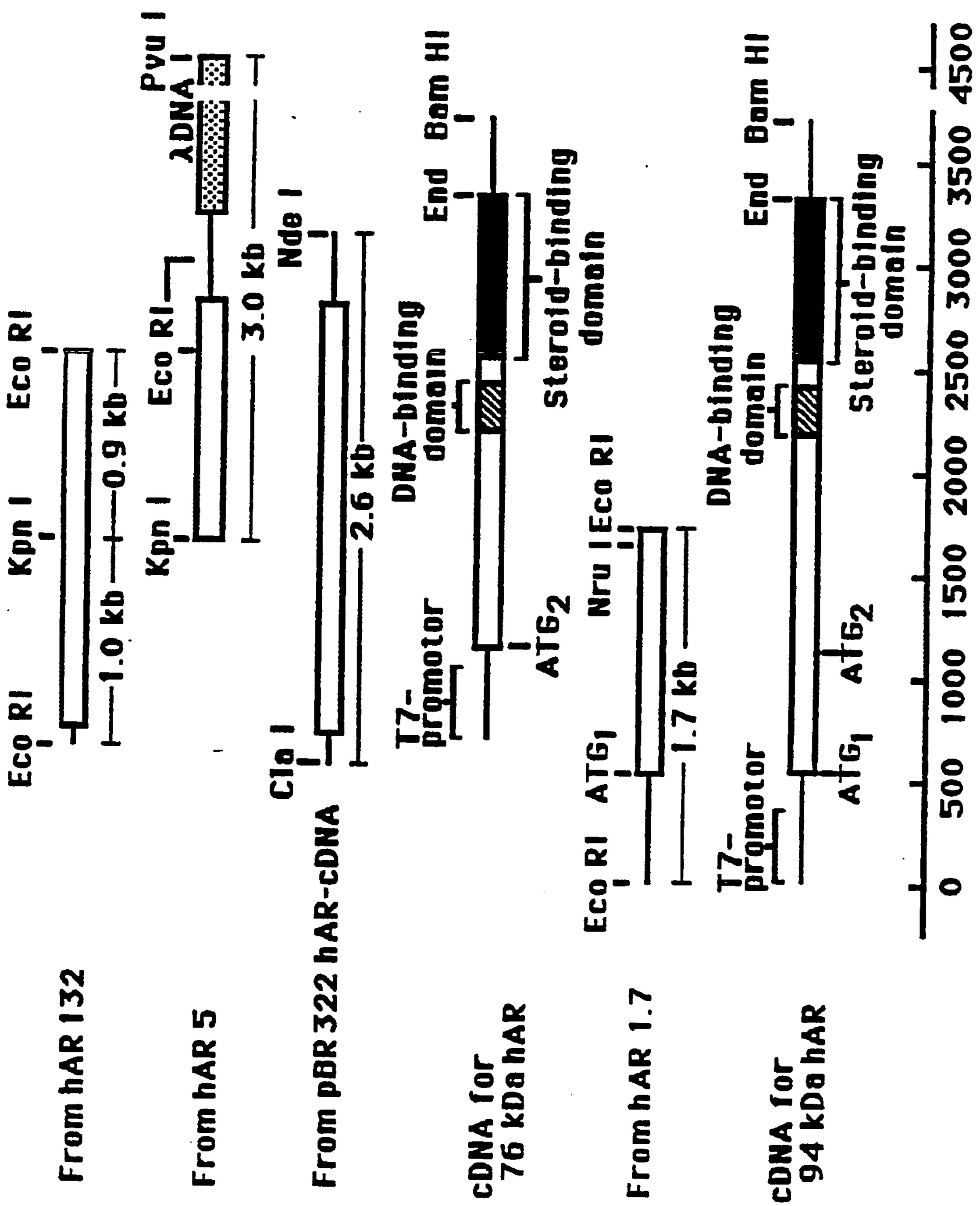
From hAR 132
Eco RI
Kpn I
Eco RI
1.0 kb
0.9 kb
From hAR 5
Kpn I
Eco RI
λDNA
Pvu I
3.0 kb
From pBR 322 hAR-cDNA
Cla I
Nde I
2.6 kb
cDNA for 76 kDa hAR
T7-promotor
ATG_2
DNA-binding domain
Steroid-binding domain
End
Bam HI
From hAR 1.7
Eco RI
ATG_1
Nru I
Eco RI
1.7 kb
cDNA for 94 kDa hAR
T7-promotor
ATG_1
ATG_2
DNA-binding domain
Steroid-binding domain
End
Bam HI
0
500
1000
1500
2000
2500
3000
3500
4500
bp

FIGURE 2

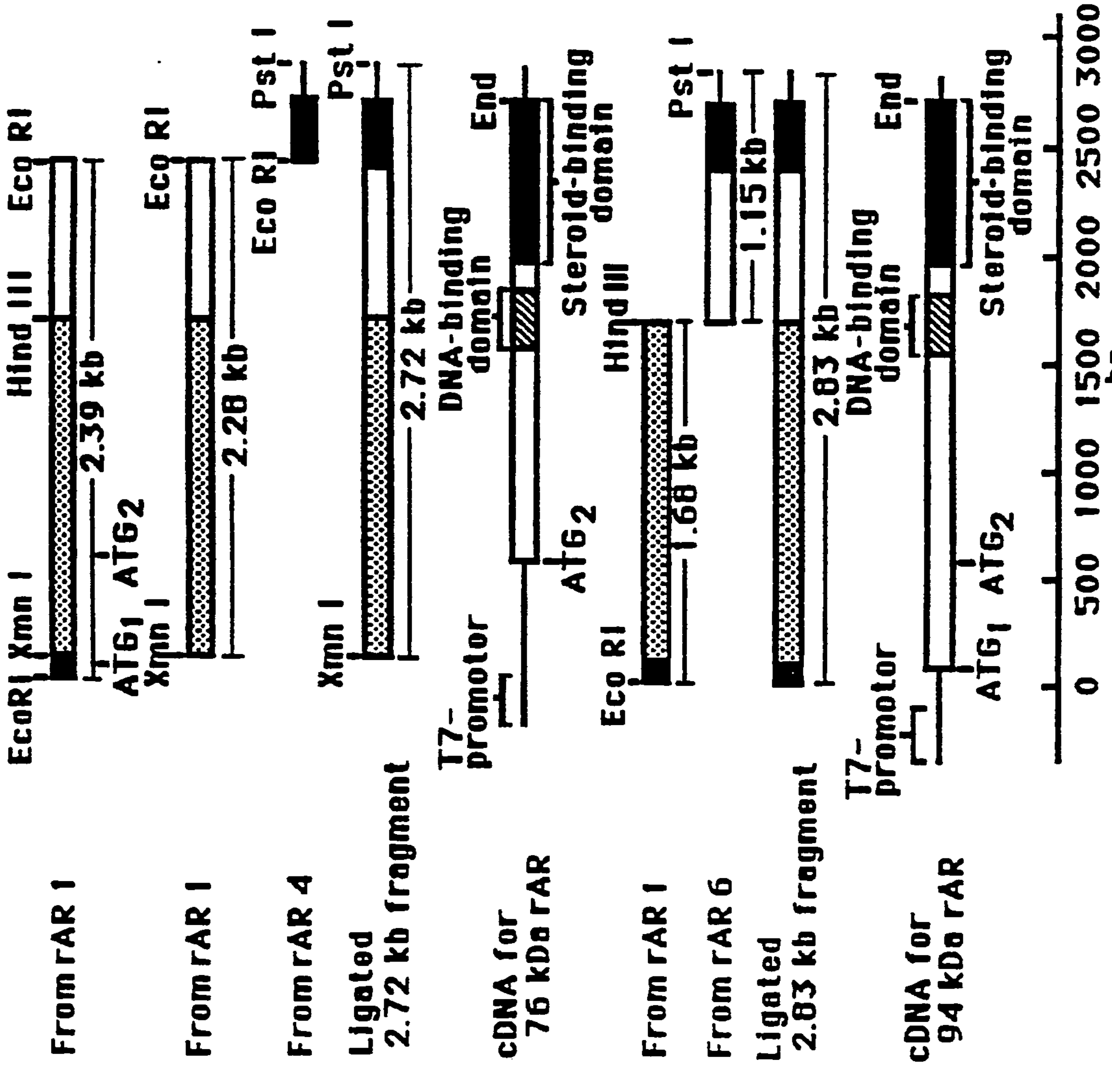
From rAR 1
From rAR 1
From rAR 4
Ligated
2.72 kb fragment
cDNA for
76 kDa rAR
From rAR 1
From rAR 6
Ligated
2.83 kb fragment
cDNA for
94 kDa rAR
EcoRI
Xmn I
ATG$_1$ ATG$_2$
Hind III
Eco RI
2.39 kb
2.28 kb
Pst I
2.72 kb
T7-
promotor
DNA-binding
domain
Steroid-binding
domain
End
1.68 kb
1.15 kb
2.83 kb
0
500
1000
1500
2000
2500
3000
bp

FIGURE 3A

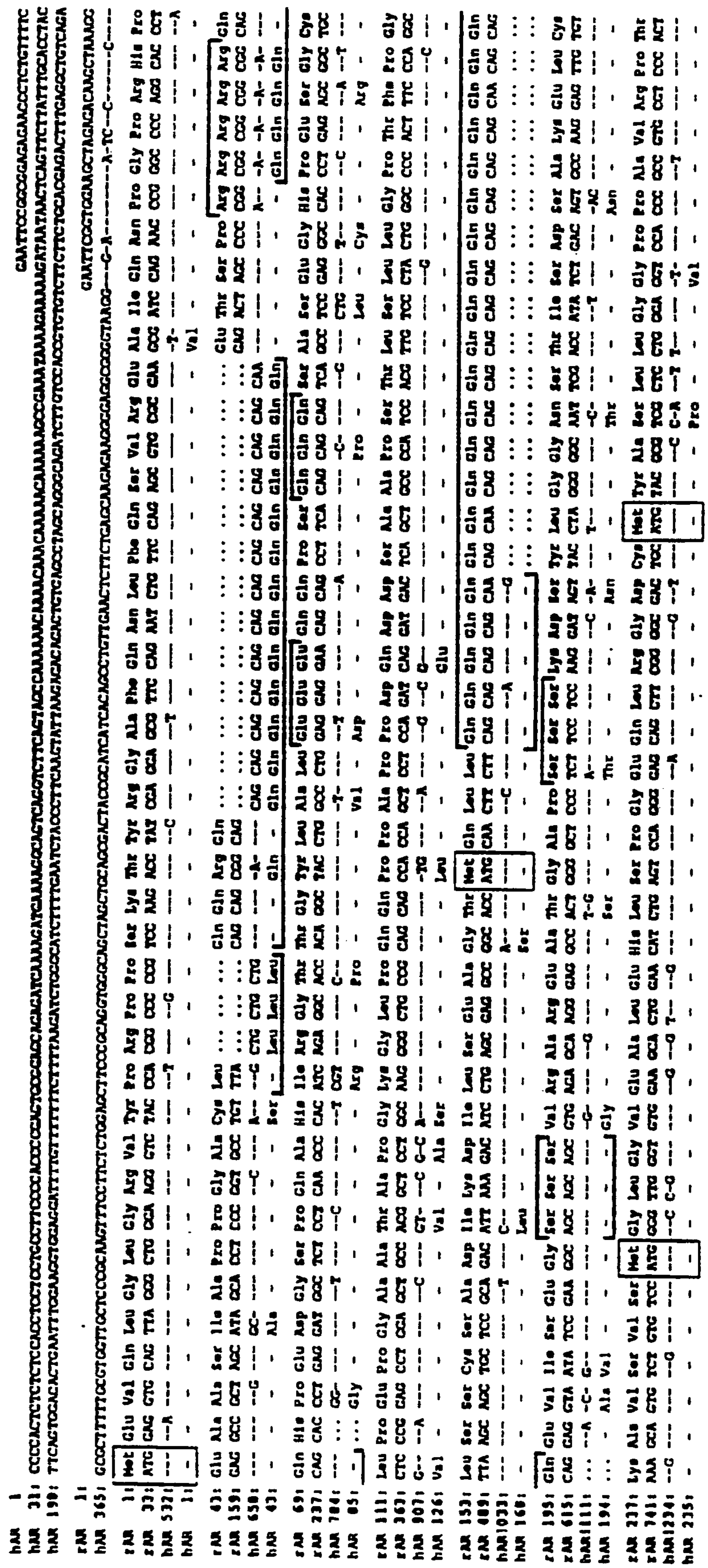

FIGURE 3B

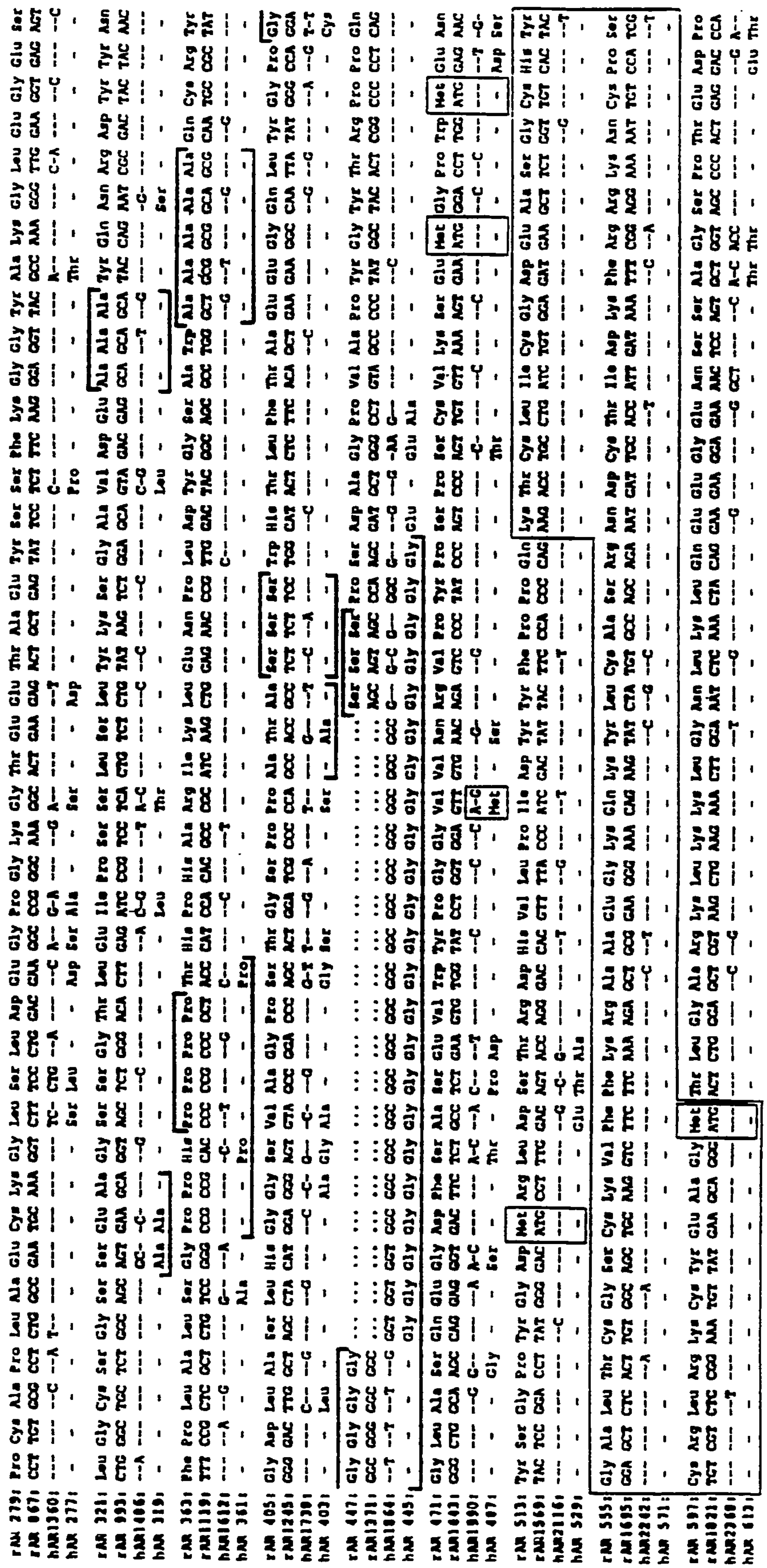

FIGURE 3C

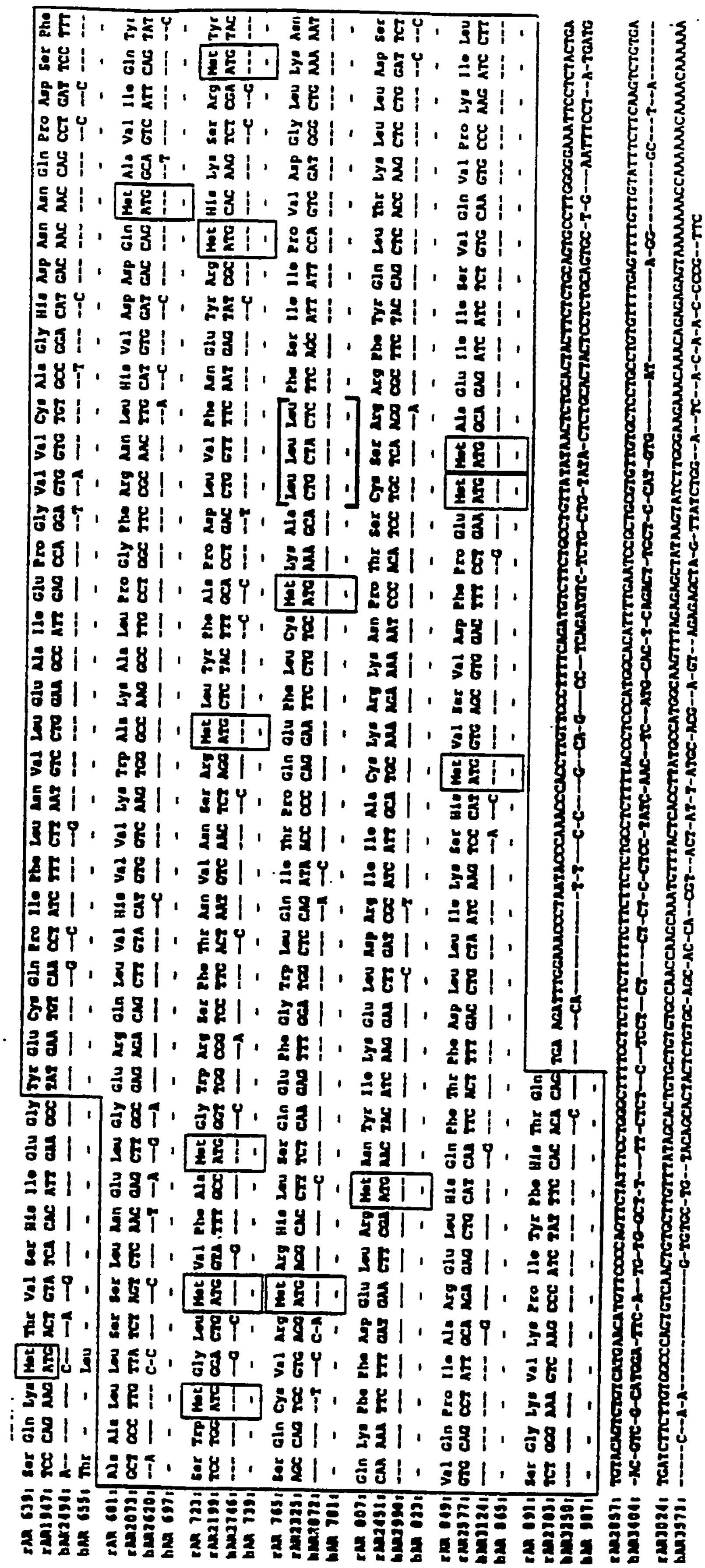

FIGURE 4A

```
                                                                                                                  GAATTCG 7
GGCCCGTCGGCTTTCTTCAACCCTCTCTTCCCGGAGCGCCCCCAATCCACGAGTGGCAGCCGCGGGACTGTCGCGTCGGCGCCCGACGCGGAGTCAGCAGGGGCGAAAAGCGGTAGATC 126
     ATG GCA ACC ATA GAA GAA ATT GCA CAT CAA ATT ATT GAA CAA CAG ATG GGA GAG ATT GTT ACA GAG CAG CAA ACT GGG CAG AAA ATC CAG 216
  1: Met Ala Thr Ile Glu Glu Ile Ala His Gln Ile Ile Glu Gln Gln Met Gly Glu Ile Val Thr Glu Gln Gln Thr Gly Gln Lys Ile Gln

     ATT GTG ACA GCA CTT GAT CAT AAT ACC CAA GGC AAG CAG TTC ATT CTG ACA AAT CAC GAC GGC TCT ACT CCA AGC AAA GTC ATT CTG GCC 306
 31: Ile Val Thr Ala Leu Asp His Asn Thr Gln Gly Lys Gln Phe Ile Leu Thr Asn His Asp Gly Ser Thr Pro Ser Lys Val Ile Leu Ala

     AGG CAA GAT TCC ACT CCG GGA AAA GTT TTC CTT ACA ACT CCA GAT GCA GCA GGT GTC AAC CAG TTA TTT TTT ACC ACT CCT GAT CTG TCT 396
 61: Arg Gln Asp Ser Thr Pro Gly Lys Val Phe Leu Thr Thr Pro Asp Ala Ala Gly Val Asn Gln Leu Phe Phe Thr Thr Pro Asp Leu Ser

     GCA CAA CAC CTG CAG CTC CTA ACA GAT AAT TCT CCA GAC CAA GGA CCA AAT AAG GTT TTT GAT CTT TGC GTA GTA TGT GGA GAC AAA GCA 486
 91: Ala Gln His Leu Gln Leu Leu Thr Asp Asn Ser Pro Asp Gln Gly Pro Asn Lys Val Phe Asp Leu Cys Val Val Cys Gly Asp Lys Ala

     TCA GGA CGT CAT TAT GGA GCA GTA ACT TGT GAA GGC TGC AAA GGA TTT TTT AAA AGA AGC ATC CGA AAA AAT TTA GTA TAT TCA TGT CGA 576
121: Ser Gly Arg His Tyr Gly Ala Val Thr Cys Glu Gly Cys Lys Gly Phe Phe Lys Arg Ser Ile Arg Lys Asn Leu Val Tyr Ser Cys Arg

     GGA TCA AAG GAT TGT ATT ATT AAT AAG CAC CAC CGA AAC CGC TGT CAA TAC TGC AGG TTA CAG AGA TGT ATT GCG TTT GGA ATG AAG CAA 666
151: Gly Ser Lys Asp Cys Ile Ile Asn Lys His His Arg Asn Arg Cys Gln Tyr Cys Arg Leu Gln Arg Cys Ile Ala Phe Gly Met Lys Gln

     GAC TCT GTC CAA TGT GAA AGA AAA CCC ATT GAA GTA TCA CGA GAA AAA TCT TCC AAC TGT GCC GCT TCA ACA GAA AAA ATC TAT ATC CGA 756
181: Asp Ser Val Gln Cys Glu Arg Lys Pro Ile Glu Val Ser Arg Glu Lys Ser Ser Asn Cys Ala Ala Ser Thr Glu Lys Ile Tyr Ile Arg

     AAG GAC CTT CGT AGC CCA TTA ACT GCA ACT CCA ACT TTT GTA ACA GAT AGT GAA AGT ACA AGG TCA ACA GGA CTG TTA GAT TCA GGA ATG 846
211: Lys Asp Leu Arg Ser Pro Leu Thr Ala Thr Pro Thr Phe Val Thr Asp Ser Glu Ser Thr Arg Ser Thr Gly Leu Leu Asp Ser Gly Met

     TTC ATG AAT ATT CAT CCA TCT GGA GTA AAA ACT GAG TCA GCT GTG CTG ATG ACA TCA GAT AAG GCT GAA TCA TGT CAG GGA GAT TTA AGT 936
241: Phe Met Asn Ile His Pro Ser Gly Val Lys Thr Glu Ser Ala Val Leu Met Thr Ser Asp Lys Ala Glu Ser Cys Gln Gly Asp Leu Ser

     ACA TTG GCC AAT GTG GTT ACA TCA TTA GCG AAT CTT GGA AAA ACT AAA GAT CTT TCT CAA AAT AGT AAT GAA ATG TCT ATG ATT GAA AGC 1026
271: Thr Leu Ala Asn Val Val Thr Ser Leu Ala Asn Leu Gly Lys Thr Lys Asp Leu Ser Gln Asn Ser Asn Glu Met Ser Met Ile Glu Ser

     TTA AGC AAT GAT GAT ACC TCT TTG TGT GAA TTT CAA GAA ATG CAG ACC AAC GGT GAT GTT TCA AGG GCA TTT GAC ACT CTT GCA AAA GCA 1116
301: Leu Ser Asn Asp Asp Thr Ser Leu Cys Glu Phe Gln Glu Met Gln Thr Asn Gly Asp Val Ser Arg Ala Phe Asp Thr Leu Ala Lys Ala
```

FIGURE 4B

```
     TTG AAT CCT GGA GAG AGC ACA GCC TGC CAG AGC TCA GTA GCG GGC ATG GAA GGA AGT GTA CAC CTA ATC ACT GGA GAT TCA AGC ATA AAT 1206
331: Leu Asn Pro Gly Glu Ser Thr Ala Cys Gln Ser Ser Val Ala Gly Met Glu Gly Ser Val His Leu Ile Thr Gly Asp Ser Ser Ile Asn
     TAC ACC GAA AAA GAG GGG CCA CTT CTC AGC GAT TCA CAT GTA GCT TTC AGG CTC ACC ATG CCT TCT CCT ATG CCT GAG TAC CTG AAT GTG 1296
361: Tyr Thr Glu Lys Glu Gly Pro Leu Leu Ser Asp Ser His Val Ala Phe Arg Leu Thr Met Pro Ser Pro Met Pro Glu Tyr Leu Asn Val
     CAC TAC ATT GGG GAG TCT GCC TCC AGA CTG CTG TTC TTA TCA ATG CAC TGG GCA CTT TCG ATT CCT TCT TTC CAG GCT CTA GGG CAA GAA 1386
391: His Tyr Ile Gly Glu Ser Ala Ser Arg Leu Leu Phe Leu Ser Met His Trp Ala Leu Ser Ile Pro Ser Phe Gln Ala Leu Gly Gln Glu
     AAC AGC ATA TCA CTG GTG AAA GCT TAC TGG AAT GAA CTT TTT ACT CTT GGT CTT GCC CAG TGC TGG CAA GTG ATG AAT GTA GCA ACT ATA 1476
421: Asn Ser Ile Ser Leu Val Lys Ala Tyr Trp Asn Glu Leu Phe Thr Leu Gly Leu Ala Gln Cys Trp Gln Val Met Asn Val Ala Thr Ile
     TTA GCA ACA TTT GTC AAT TGT CTT CAC AAT AGT CTT CAA CAA GAT GCC AAG GTA ATT GCA GCC CTC ATT CAT TTC ACA AGA CGA GCA ATC 1566
451: Leu Ala Thr Phe Val Asn Cys Leu His Asn Ser Leu Gln Gln Asp Ala Lys Val Ile Ala Ala Leu Ile His Phe Thr Arg Arg Ala Ile
     ACT GAT TTA TAA ATGCTTAACTATAGAATGGCTTATGACTACCCAAAACAGTGCCCCATCAACAAATGGGGAAAATTGCCTTTTGAGCTCAGGAATAATTTATAAATTGGGGACT 1681
481: Thr Asp Leu ***
     ACCTTTTAGTTCTTTAGCATATTCTATTTCTTATTGTTTTATATAATTTTTAAATCATTTGCTTCCTCCTTATGTTTAACAGCAGAGGGGTAATCACCTTAAAATGTCATCAAAAATAG 1800
     ATCTACTAGAAGGCAGCATCACATTCCCATCTTACTTATGGACTCCTACCCCTGGTTCATGTCTTATATGCCTGTAATGGTTATAAAGCCTACCTTCAGGAAAGCTATGGTTGACTAAT 1919
     TACTAATGGATGGGTTTTAAACATGTCCCTCTACAATAAATTAAAATCTTTCAATGTTTGAATATAATGTGGAGGTGTTTACCTGAGGGCCTCTCTATCTCCCCGAATTC     2029
```

*: 11 of 30 TR2 clones have extra 429 bp insert here which create a termination codon TAG.

```
GTATGTATTAGCTTTTAAGGAGAAAATACTTTTTAAAGATTCCAGCAAACTACAAGAGTATTGAAATTAACAAAATATGTCAAATATGTATACTTTTTAGTTTACAGTTTTCCAACTA
AAATATAAGAATACAGACCTACGTATGTACTTTTTTATTTATTCAACTAATATTTATTGAATACCTGTGTAAAAAGCACGTATTTGTAGTTTGTGTTTGGCAAAAAGAATGAGAGATGC
AAATTGGTTTGTTTTACTAATCTAAAGCAACTTTGTTGAACTTGCACATAATTTCTAAAGATTGATGGTTATCTTTGGAGTTTAGTATGGTAGCCATGTCTCCTATTAGCAGCATTAAG
CTTACCTACAGCTTACATTTCTAATTGTCTGTAATCCTATATTGTGATATAATAGTTTAACACATTTTTGTAG
```

Human TR2-9

```
  1:

  8:GGCCCGTCGGCTTTCTTCAACCCTCTCTTCCCGGAGCGCCCCCAATCCACGAGTGGCAG

127:ATG GCA ACC ATA GAA GAA ATT GCA CAT CAA ATT ATT GAA CAA CAG
  1:met ala thr ile glu glu ile ala his gln ile ile glu gln gln

217:ATT GTG ACA GCA CTT GAT CAT AAT ACC CAA GGC AAG CAG TTC ATT
 31:ile val thr ala leu asp his asn thr gln gly lys gln phe ile

307:AGG CAA GAT TCC ACT CCG GGA AAA GTT TTC CTT ACA ACT CCA GAT
 61:arg gln asp ser thr pro gly lys val phe leu thr thr pro asp

397:GCA CAA CAC CTG CAG CTC CTA ACA GAT AAT TCT CCA GAC CAA GGA
 91:ala gln his leu gln leu leu thr asp asn ser pro asp gln gly

487:TCA GGA CGT CAT TAT GGA GCA GTA ACT TGT GAA GGC TGC AAA GGA
121:ser gly arg his tyr gly ala val thr cys glu gly cys lys gly

577:GGA TCA AAG GAT TGT ATT ATT AAT AAG CAC CAC CGA AAC CGC TGT
151:gly ser lys asp cys ile ile asn lys his his arg asn arg cys

667:GAC TCT GTC CAA TGT GAA AGA AAA CCC ATT GAA GTA TCA CGA GAA
181:asp ser val gln cys glu arg lys pro ile glu val ser arg glu

757:AAG GAC CTT CGT AGC CCA TTA ACT GCA ACT CCA ACT TTT GTA ACA
211:lys asp leu arg ser pro leu thr ala thr pro thr phe val thr
```

FIGURE 5
TOP LEFT SIDE

```
                                                         GAATTCG

CCGCGGGACTGTCGCGTCGGCGCCCGACGCGGAGTCAGCAGGGGCGAAAAGCGGTAGATC

ATG GGA GAG ATT GTT ACA GAG CAG CAA ACT GGG CAG AAA ATC CAG
met gly glu ile val thr glu gln gln thr gly gln lys ile gln

CTG ACA AAT CAC GAC GGC TCT ACT CCA AGC AAA GTC ATT CTG GCC
leu thr asn his asp gly ser thr pro ser lys val ile leu ala

GCA GCA GGT GTC AAC CAG TTA TTT TTT ACC ACT CCT GAT CTG TCT
ala ala gly val asn gln leu phe phe thr thr pro asp leu ser

CCA AAT AAG GTT TTT GAT CTT TGC GTA GTA TGT GGA GAC AAA GCA
pro asn lys val phe asp leu cys val val cys gly asp lys ala

TTT TTT AAA AGA AGC ATC CGA AAA AAT TTA GTA TAT TCA TGT CGA
phe phe lys arg ser ile arg lys asn leu val tyr ser cys arg

CAA TAC TGC AGG TTA CAG AGA TGT ATT GCG TTT GGA ATG AAG CAA
gln tyr cys arg leu gln arg cys ile ala phe gly met lys gln

AAA TCT TCC AAC TGT GCC GCT TCA ACA GAA AAA ATC TAT ATC CGA
lys ser ser asn cys ala ala ser thr glu lys ile tyr ile arg

GAT AGT GAA AGT ACA AGG TCA ACA GGA CTG TTA GAT TCA GGA ATG
asp ser glu ser thr arg ser thr gly leu leu asp ser gly met
```

FIGURE 5
TOP RIGHT SIDE

```
847:TTC ATG AAT ATT CAT CCA TCT GGA GTA AAA ACT GAG TCA GCT
241:phe met asn ile his pro ser gly val lys thr glu ser ala

937:ACA TTG GCC AAT GTG GTT ACA TCA TTA GCG AAT CTT GGA AAA
271:thr leu ala asn val val thr ser leu ala asn leu gly lys

1027:TTA AGC AAT GAT GAT ACC TCT TTG TGT GAA TTT CAA GAA ATG
301:leu ser asn asp asp thr ser leu cys glu phe gln glu met

1117:TTG AAT CCT GGA GAG AGC ACA GCC TGC CAG AGC TCA GTA GCG
331:leu asn pro gly glu ser thr ala cys gln ser ser val ala

1207:TAC ACC GAA AAA GAG GGG CCA CTT CTC AGC GAT TCA CAT GTA
361:tyr thr glu lys glu gly pro leu leu ser asp ser his val

1297:CAC TAC ATT GGG GAG TCT GCC TCC AGA CTG CTG TTC TTA TCA
391:his tyr ile gly glu ser ala ser arg leu leu phe leu ser

1387:AAC AGC ATA TCA CTG GTG AAA GCT TAC TGG AAT GAA CTT TTT
421:asn ser ile ser leu val lys ala tyr trp asn glu leu phe

1477:TTA GCA ACA TTT GTC AAT TGT CTT CAC AAT AGT CTT CAA CAA
451:leu ala thr phe val asn cys leu his asn ser leu gln gln

1570:CATTCCCATCTTACTTATGGACTCCTACCCCTGGTTCATGTCTTATATGCCTGTA

1697:CATGTCCCTCTACAATAAATTAAAATCTTTCAATGTTTGAATATAATGTGGAGGT
```

FIGURE 5
BOTTOM LEFT SIDE

GTG CTG ATG ACA TCA GAT AAG GCT GAA TCA TGT CAG GGA GAT TTA AGT
val leu met thr ser asp lys ala glu ser cys gln gly asp leu ser

ACT AAA GAT CTT TCT CAA AAT AGT AAT GAA ATG TCT ATG ATT GAA AGC
thr lys asp leu ser gln asn ser asn glu met ser met ile glu ser

CAG ACC AAC GGT GAT GTT TCA AGG GCA TTT GAC ACT CTT GCA AAA GCA
gln thr asn gly asp val ser arg ala phe asp thr leu ala lys ala

GGC ATG GAA GGA AGT GTA CAC CTA ATC ACT GGA GAT TCA AGC ATA AAT
gly met glu gly ser val his leu ile thr gly asp ser ser ile asn

GCT TTC AGG CTC ACC ATG CCT TCT CCT ATG CCT GAG TAC CTG AAT GTG
ala phe arg leu thr met pro ser pro met pro glu tyr leu asn val

ATG CAC TGG GCA CTT TCG ATT CCT TCT TTC CAG GCT CTA GGG CAA GAA
met his trp ala leu ser ile pro ser phe gln ala leu gly gln glu

ACT CTT GGT CTT GCC CAG TGC TGG CAA GTG ATG AAT GTA GCA ACT ATA
thr leu gly leu ala gln cys trp gln val met asn val ala thr ile

GCA GAG GGG TAA TCACCTTAAAATGTCATCAAAAATAGATCTACTAGAAGGCAGCATCA
ala glu gly STOP

ATGGTTATAAAGCCTACCTTCAGGAAAGCTATGGTTGACTAATTACTAATGGATGGGTTTTAAA

GTTTACCTGAGGGCCTCTCTATCTCCCCGAATTC

FIGURE 5
BOTTOM RIGHT SIDE

Human TRα-11

```
  1:

 57:ATG GCA ACC ATA GAA GAA ATT GCA CAT CAA ATT ATT GAA
  1:met ala thr ile glu glu ile ala his gln ile ile glu

147:ATT GTG ACA GCA CTT GAT CAT AAT ACC CAA GGC AAG CAG
 31:ile val thr ala leu asp his asn thr gln gly lys gln

237:AGG CAA GAT TCC ACT CCG GGA AAA GTT TTC CTT ACA ACT
 61:arg gln asp ser thr pro gly lys val phe leu thr thr

327:GCA CAA CAC CTG CAG CTC CTA ACA GAT AAT TCT CCA GAC
 91:ala gln his leu gln leu leu thr asp asn ser pro asp

417:TCA GGA CGT CAT TAT GGA GCA GTA ACT TGT GAA GGC TGC
121:ser gly arg his tyr gly ala val thr cys glu gly cys

507:GGA TCA AAG GAT TGT ATT ATT AAT AAG CAC CAC CGA AAC
151:gly ser lys asp cys ile ile asn lys his his arg asn

597:GAC TCT GTC CAA TGT GAA AGA AAA CCC ATT GAA GTA TCA
181:asp ser val gln cys glu arg lys pro ile glu val ser

687:AAG GAC CTT CGT AGC CCA TTA ACT GCA ACT CCA ACT TTT
211:lys asp leu arg ser pro leu thr ala thr pro thr phe

777:TTC ATG AAT ATT CAT CCA TCT GGA GTA AAA ACT GAG TCA
241:phe met asn ile his pro ser gly val lys thr glu ser

867:ACA TTG GCC AAT GTG GTT ACA TCA TTA GCG AAT CTT GGA
271:thr leu ala asn val val thr ser leu ala asn leu gly

957:TTA AGC AAT GAT GAT ACC TCT TTG TGT GAA TTT CAA GAA
301:leu ser asn asp asp thr ser leu cys glu phe gln glu
```

FIGURE 6
TOP LEFT SIDE

GGGACTGTCGCGTCGGCCCCCGACGCGGAGTCAGCAGGGGCGAAAAGCGGTAGATC

CAA CAG ATG GGA GAG ATT GTT ACA GAG CAG CAA ACT GGG CAG AAA ATC CAG
gln gln met gly glu ile val thr glu gln gln thr gly gln lys ile gln

TTC ATT CTG ACA AAT CAC GAC GGC TCT ACT CCA AGC AAA GTC ATT CTG GCC
phe ile leu thr asn his asp gly ser thr pro ser lys val ile leu ala

CCA GAT GCA GCA GGT GTC AAC CAG TTA TTT TTT ACC ACT CCT GAT CTG TCT
pro asp ala ala gly val asn gln leu phe phe thr thr pro asp leu ser

CAA GGA CCA AAT AAG GTT TTT GAT CTT TGC GTA GTA TGT GGA GAC AAA GCA
gln gly pro asn lys val phe asp leu cys val val cys gly asp lys ala

AAA GGA TTT TTT AAA AGA AGC ATC CGA AAA AAT TTA GTA TAT TCA TGT CGA
lys gly phe phe lys arg ser ile arg lys asn leu val tyr ser cys arg

CGC TGT CAA TAC TGC AGG TTA CAG AGA TGT ATT GCG TTT GGA ATG AAG CAA
arg cys gln tyr cys arg leu gln arg cys ile ala phe gly met lys gln

CGA GAA AAA TCT TCC AAC TGT GCC GCT TCA ACA GAA AAA ATC TAT ATC CGA
arg glu lys ser ser asn cys ala ala ser thr glu lys ile tyr ile arg

GTA ACA GAT AGT GAA AGT ACA AGG TCA ACA GGA CTG TTA GAT TCA GGA ATG
val thr asp ser glu ser thr arg ser thr gly leu leu asp ser gly met

GCT GTG CTG ATG ACA TCA GAT AAG GCT GAA TCA TGT CAG GGA GAT TTA AGT
ala val leu met thr ser asp lys ala glu ser cys gln gly asp leu ser

AAA ACT AAA GAT CTT TCT CAA AAT AGT AAT GAA ATG TCT ATG ATT GAA AGC
lys thr lys asp leu ser gln asn ser asn glu met ser met ile glu ser

ATG CAG ACC AAC GGT GAT GTT TCA AGG GCA TTT GAC ACT CTT GCA AAA GCA
met gln thr asn gly asp val ser arg ala phe asp thr leu ala lys ala

FIGURE 6
TOP RIGHT SIDE

```
1047:TTG AAT CCT GGA GAG AGC ACA GCC TGC CAG AGC TCA GTA GCC
 331:leu asn pro gly glu ser thr ala cys gln ser ser val ala

1137:TAC ACC GAA AAA GAG GGG CCA CTT CTC AGC GAT TCA CAT GTA
 361:tyr thr glu lys glu gly pro leu leu ser asp ser his val

1227:CAC TAC ATT GGG GAG TCT GCC TCC AGA CTG CTG TTC TTA TCA
 391:his tyr ile gly glu ser ala ser arg leu leu phe leu ser

1317:AAC AGC ATA TCA CTG GTG AAA GCT TAC TGG AAT GAA CTT TTT
 421:asn ser ile ser leu val lys ala tyr trp asn glu leu phe

1407:TTA GCA ACA TTT GTC AAT TGT CTT CAC AAT AGT CTT CAA CAA
 451:leu ala thr phe val asn cys leu his asn ser leu gln gln

1497:AAA CTA CAG GAG TTT TGT AAC AGC ATG GTT AAA CTC TGC ATT
 481:lys leu gln glu phe cys asn ser met val lys leu cys ile

1587:GAT CAT CCA AGC CTA GAA AAC ATG GAA CTG ATA GAG AAA TTT
 511:asp his pro ser leu glu asn met glu leu ile glu lys phe

1677:CCA GAT GAC ACC TAC AGG TTA TCC AGA CTA CTA CTC AGA TTG
 541:pro asp asp thr tyr arg leu ser arg leu leu leu arg leu

1767:AAA GGT CTC ATT GGC AAT ATA CGA ATT GAC AGT GTT ATC CCA
 571:lys gly leu ile gly asn ile arg ile asp ser val ile pro

1857:CAC AGC ATT TGA AAACTGTGACTGCAGTGCTGTAAACTTAACTGTTCTTTG
 601:his ser ile STOP

1972:AAGTAACCAGAATCCAAGGTATTTTTATTTTAGCTTCCCTTAAGAATTTTTGAAG

2091:CACTACAAATTTATTCTTGGTGAAGATGATACCTGAAGCTGTCACCTCTTGATTA

2210:AAAAAAAAACCC
```

FIGURE 6
BOTTOM LEFT SIDE

GGC ATG GAA GGA AGT GTA CAC CTA ATC ACT GGA GAT TCA AGC ATA AAT

gly met glu gly ser val his leu ile thr gly asp ser ser ile asn

GCT TTC AGG CTC ACC ATG CCT TCT CCT ATG CCT GAG TAC CTG AAT GTG

ala phe arg leu thr met pro ser pro met pro glu tyr leu asn val

ATG CAC TGG GCA CTT TCG ATT CCT TCT TTC CAG GCT CTA GGG CAA GAA

met his trp ala leu ser ile pro ser phe gln ala leu gly gln glu

ACT CTT GGT CTT GCC CAG TGC TGG CAA GTG ATG AAT GTA GCA ACT ATA

thr leu gly leu ala gln cys trp gln val met asn val ala thr ile

GAT AAA ATG TCA ACA GAA AGA AGA AAA TTA TTG ATG GAG CAC ATC TTC

asp lys met ser thr glu arg arg lys leu leu met glu his ile phe

GAT GGA TAC GAA TAT GCC TAC CTG AAG GCA ATA GTA CTC TTC AGT CCA

asp gly tyr glu tyr ala tyr leu lys ala ile val leu phe ser pro

CAG GAA AAG GCT TAT GTG GAA TTC CAA GAT TAT ATA ACC AAA ACA TAT

gln glu lys ala tyr val glu phe gln asp tyr ile thr lys thr tyr

CCA GCT TTA AGA CTG ATG AAT GCT ACC ATC ACT GAA GAA TTG TTT TTC

pro ala leu arg leu met asn ala thr ile thr glu glu leu phe phe

CAT ATT TTG AAA ATG GAG CCT GCA GAT TAT AAC TCT CAA ATA ATT GGT

his ile leu lys met glu pro ala asp tyr asn ser gln ile ile gly

CCAGAACACAAGACACCAAATTGAACTCACTGCTTTTGAGGCATCTGGAAATTTTTACTTTAAA

TGACTGGGCAGGCAGCAGAAATTAAATGAATTTTTCTTCCTGATTCCTTTAAATGAATATGAAA

TCTAAACTAAGCGCTCATTCTATTTTATAAAACA<u>AATAAA</u>TTAGTCTCTTTTTTCTGAAAAAAA

FIGURE 6
BOTTOM RIGHT SI

FIGURE 7

h-GR 419	Lys	Leu	Cys	Leu	Val	Cys	Ser	Asp	Glu	Ala	Ser	Gly	Cys	His	Tyr	Gly	Val	Leu	Thr	Cys	Gly	Ser	Cys	Lys
h-MR 601	Lys	Ile	Cys	Leu	Val	Cys	Gly	Asp	Glu	Ala	Ser	Gly	Cys	His	Tyr	Gly	Val	Val	Thr	Cys	Gly	Ser	Cys	Lys
h-PR 565	Lys	Ile	Cys	Leu	Ile	Cys	Gly	Asp	Glu	Ala	Ser	Gly	Cys	His	Tyr	Gly	Val	Leu	Thr	Cys	Gly	Ser	Cys	Lys
h-AR	Lys	Thr	Cys	Leu	Ile	Cys	Gly	Asp	Glu	Ala	Ser	Gly	Cys	His	Tyr	Gly	Ala	Leu	Thr	Cys	Gly	Ser	Cys	Lys
r-AR	Lys	Thr	Cys	Leu	Ile	Cys	Gly	Asp	Glu	Ala	Ser	Gly	Cys	His	Tyr	Gly	Ala	Leu	Thr	Cys	Gly	Ser	Cys	Lys
h-ER 183	Arg	Tyr	Cys	Ala	Val	Cys	Asn	Asp	Tyr	Ala	Ser	Gly	Tyr	His	Tyr	Gly	Val	Trp	Ser	Cys	Glu	Gly	Cys	Lys
h-TR2	Asp	Leu	Cys	Val	Val	Cys	Gly	Asp	Lys	Ala	Ser	Gly	Arg	His	Tyr	Gly	Ala	Val	Thr	Cys	Glu	Gly	Cys	Lys
v-erbA35	Glu	Gln	Cys	Val	Val	Cys	Gly	Asp	Lys	Ala	Thr	Gly	Tyr	His	Tyr	Arg	Cys	Ile	Thr	Cys	Glu	Gly	Cys	Lys
c-VDR	Arg	Ile	Cys	Gly	Val	Cys	Gly	Asp	Arg	Ala	Thr	Gly	Phe	His	Phe	Asn	Ala	Met	Thr	Cys	Glu	Gly	Cys	Lys

h-GR 439	Val	Phe	Phe	Lys	Arg	Ala	Val	Glu	Gly	Gln	His	Asn	Tyr	Leu	Cys	Ala	Gly	Arg	Asn	Asp	Cys	Ile	Ile	Asp
h-MR 625	Val	Phe	Phe	Lys	Arg	Ala	Val	Glu	Gly	Gln	His	Asn	Tyr	Leu	Cys	Ala	Gly	Arg	Asn	Asp	Cys	Ile	Ile	Asp
h-PR 589	Val	Phe	Phe	Lys	Arg	Ala	Met	Glu	Gly	Gln	His	Asn	Tyr	Leu	Cys	Ala	Gly	Arg	Asn	Asp	Cys	Ile	Val	Asp
h-AR	Val	Phe	Phe	Lys	Arg	Ala	Ala	Glu	Gly	Lys	Gln	Lys	Tyr	Leu	Cys	Ala	Ser	Arg	Asn	Asp	Cys	Thr	Ile	Asp
r-AR	Val	Phe	Phe	Lys	Arg	Ala	Ala	Glu	Gly	Lys	Gln	Lys	Tyr	Leu	Cys	Ala	Ser	Arg	Asn	Asp	Cys	Thr	Ile	Asp
h-ER 207	Ala	Phe	Phe	Lys	Arg	Ser	Ile	Gln	Gly	His	Asn	Asp	Tyr	Met	Cys	Pro	Ala	Thr	Asn	Gln	Cys	Thr	Ile	Asp
h-TR2	Gly	Phe	Phe	Lys	Arg	Ser	Ile	Arg	Lys	Asn	Leu	Val	Tyr	Ser	Cys	Arg	Gly	Ser	Lys	Asp	Cys	Ile	Ile	Asn
v-erbA59	Ser	Phe	Phe	Arg	Arg	Thr	Ile	Gln	Lys*	His	Pro	Thr	Tyr	Ser	Cys	Thr	Tyr	Asp	Gly	Cys	Cys	Val	Ile	Asp
c-VDR	Gly	Phe	Phe	Arg	Arg	Ser	Met	Lys	Arg	Lys	Ala	Met	Phe	Thr	Cys	Pro	Phe	Asn	Gly	Asp	Cys	Lys	Ile	Thr

h-GR 463	Lys	Ile	Arg	Arg	Lys	Asn	Cys	Pro	Ala	Cys	Arg	Tyr	Arg	Lys	Cys	Leu	Gln	Ala	Gly	Met	Asn	Leu	Glu	Ala
h-MR 649	Lys	Ile	Arg	Arg	Lys	Asn	Cys	Pro	Ala	Cys	Arg	Leu	Gln	Lys	Cys	Leu	Gln	Ala	Gly	Met	Asn	Leu	Gly	Ala
h-PR 613	Lys	Ile	Arg	Arg	Lys	Asn	Cys	Pro	Ala	Cys	Arg	Leu	Arg	Lys	Cys	Cys	Gln	Ala	Gly	Met	Val	Leu	Gly	Gly
h-AR	Lys	Phe	Arg	Arg	Lys	Asn	Cys	Pro	Ser	Cys	Arg	Leu	Arg	Lys	Cys	Tyr	Glu	Ala	Gly	Met	Thr	Leu	Gly	Ala
r-AR	Lys	Phe	Arg	Arg	Lys	Asn	Cys	Pro	Ser	Cys	Arg	Leu	Arg	Lys	Cys	Tyr	Glu	Ala	Gly	Met	Thr	Leu	Gly	Ala
h-ER 231	Lys	Asn	Arg	Arg	Lys	Ser	Cys	Gln	Ala	Cys	Arg	Leu	Arg	Lys	Cys	Tyr	Glu	Val	Gly	Met	Met	Lys	Gly	Gly
h-TR2	Lys	His	His	Arg	Asn	Arg	Cys	Gln	Tyr	Cys	Arg	Leu	Gln	Arg	Cys	Ile	Ala	Phe	Gly	Met	Lys	Gln	Asp	Cys
v-erbA85	Lys	Ile	Thr	Arg	Asn	Gln	Cys	Gln	Leu	Cys	Arg	Phe	Lys	Lys	Cys	Ile	Ser	Val	Gly	Met	Ala	Met	Asp	Leu
c-VDR	Lys	Asp	Asn	Arg	Arg	His	Cys	Gln	Ala	Cys	Arg	Leu	Lys	Arg	Cys	Val	Asp	Ile	Gly	Met	Met	Lys		

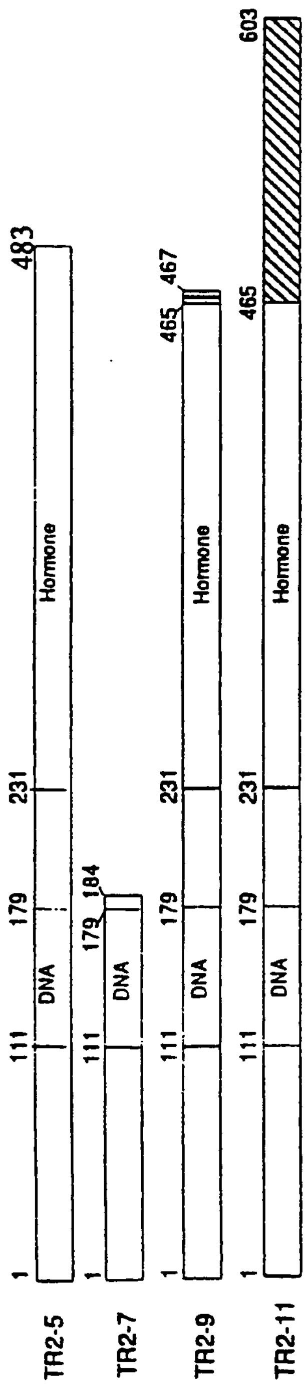
TR2-5
1
111
DNA
179
231
Hormone
483
TR2-7
1
111
DNA
179
184
TR2-9
1
111
DNA
179
231
Hormone
465
467
TR2-11
1
111
DNA
179
231
Hormone
465
603

FIGURE 8

323 amino acids from TRP E protein

```
          170         180         190         200         210         220         230         240         250
ATG CAA ACA CAA AAA CCG ACT CTC GAA CTG CTA ACC TGC GAA GGC GCT TAT CGC GAC AAT CCC ACC GCG CTT TTT CAC CAG TTG TGT GGG
MET GLN THR GLN LYS PRO THR LEU GLU LEU LEU THR CYS GLU GLY ALA TYR ARG ASP ASN PRO THR ALA LEU PHE HIS GLN LEU CYS GLY
 1                            10                            20                            30
          260         270         280         290         300         310         320         330         340
GAT CGT CCG GCA ACG CTG CTG CTG GAA TCC GCA GAT ATC GAC AGC AAA GAT GAT TTA AAA AGC CTG CTG CTG GTA GAC AGT GCG CTG CGC
ASP ARG PRO ALA THR LEU LEU LEU GLU SER ALA ASP ILE ASP SER LYS ASP ASP LEU LYS SER LEU LEU LEU VAL ASP SER ALA LEU ARG
                              40                            50                            60
          350         360         370         380         390         400         410         420         430
ATT ACA GCT TTA GGT GAC ACT GTC ACA ATC CAG GCA CTT TCC GGC AAC GGC GAA GCC CTC CTG GCA CTA CTG GAT AAC GCC CTG CCT GCG
ILE THR ALA LEU GLY ASP THR VAL THR ILE GLN ALA LEU SER GLY ASN GLY GLU ALA LEU LEU ALA LEU LEU ASP ASN ALA LEU PRO ALA
                              70                            80                            90
          440         450         460         470         480         490         500         510         520
GGT GTG GAA AGT GAA CAA TCA CCA AAC TGC CGT GTG CTG CGC TTC CCC CCT GTC AGT CCA CTG CTG GAT GAA GAC GCC CGC TTA TGC TCC
GLY VAL GLU SER GLU GLN SER PRO ASN CYS ARG VAL LEU ARG PHE PRO PRO VAL SER PRO LEU LEU ASP GLU ASP ALA ARG LEU CYS SER
                             100                           110                           120
          530         540         550         560         570         580         590         600         610
CTT TCG GTT TTT GAC GCT TTC CGT TTA TTG CAG AAT CTG TTG AAT GTA CCG AAG GAA GAA CGA GAA GCC ATG TTC TTC AGC GGC CTG TTC
LEU SER VAL PHE ASP ALA PHE ARG LEU LEU GLN ASN LEU LEU ASN VAL PRO LYS GLU GLU ARG GLU ALA MET PHE PHE SER GLY LEU PHE
                             130                           140                           150
          620         630         640         650         660         670         680         690         700
TCT TAT GAC CTT GTG GCG GGA TTT GAA GAT TTA CCG CAA CTG TCA GCG GAA AAT AAC TGC CCT GAT TTC TGT TTT TAT CTC GCT GAA ACG
SER TYR ASP LEU VAL ALA GLY PHE GLU ASP LEU PRO GLN LEU SER ALA GLU ASN ASN CYS PRO ASP PHE CYS PHE TYR LEU ALA GLU THR
                             160                           170                           180
          710         720         730         740         750         760         770         780         790
CTG ATG GTG ATT GAC CAT CAG AAA AAA AGC ACC CGT ATT CAG GCC AGC CTG TTT GCT CCG AAT GAA GAA GAA AAA CAA CGT CTG ACT GCT
LEU MET VAL ILE ASP HIS GLN LYS LYS SER THR ARG ILE GLN ALA SER LEU PHE ALA PRO ASN GLU GLU GLU LYS GLN ARG LEU THR ALA
                             190                           200                           210
          800         810         820         830         840         850         860         870         880
CGC CTG AAC GAA CTA CGT CAG CAA CTG ACC GAA GCC GCG CCG CCG CTG CCA GTG GTT TCC GTG CCG CAT ATG CGT TGT GAA TGT AAT CAG
ARG LEU ASN GLU LEU ARG GLN GLN LEU THR GLU ALA ALA PRO PRO LEU PRO VAL VAL SER VAL PRO HIS MET ARG CYS GLU CYS ASN GLN
                             220                           230                           240
          890         900         910         920         930         940         950         960         970
AGC GAT GAA GAG TTC GGT GGC GTA GTG CGT TTG TTG CAA AAA GCG ATT CGC GCT GGA GAA ATT TTC CAG GTG GTG CCA TCT CGC CGT TTC
SER ASP GLU GLU PHE GLY GLY VAL VAL ARG LEU LEU GLN LYS ALA ILE ARG ALA GLY GLU ILE PHE GLN VAL VAL PRO SER ARG ARG PHE
                             250                           260                           270
          980         990        1000        1010        1020        1030        1040        1050        1060
TCT CTG CCC TGC CCG TCA CCG CTG GCG GCC TAT TAC GTG CTG AAA AAG AGT AAT CCC AGC CCG TAC ATG TTT TTT ATG CAG GAT AAT GAT
SER LEU PRO CYS PRO SER PRO LEU ALA ALA TYR TYR VAL LEU LYS LYS SER ASN PRO SER PRO TYR MET PHE PHE MET GLN ASP ASN ASP
                             280                           290                           300
         1070        1080        1090        1100        1110        1120        1130
TTC ACC CTA TTT GGC GCG TCG CCG GAA AGC TCG CTC AAG TAT GAT GCC ACC AGC CGC CAG ATT GAG ATC
PHE THR LEU PHE GLY ALA SER PRO GLU SER SER LEU LYS TYR ASP ALA THR SER ARG GLN ILE GLU ILE
                             310                           320
```

6 amino acid linker

CCC CGG AAT TCG AGC TCG
Pro Arg Asn Ser Ser Ser

FIGURE 9A

```
GGGACACTTGAACTGCCGTCTACCCTGTCTCTCTACAAGTCCGGAGCACTGGACGAGGCAGCTGCGTACCAGAGTCGCGACTACTACAACTTT
GlyThrLeuGluLeuProSerThrLeuSerLeuTyrLysSerGlyAlaLeuAspGluAlaAlaAlaTyrGlnSerArgAspTyrTyrAsnPhe
CCACTGGCTCTGGCCGGACCGCCGCCCCCTCCGCCGCCTCCCCATCCCCACGCTCGCATCAAGCTGGAGAACCCGCTGGACTACGGCAGCGCCTGGGCGGCTGCGGCGGCGCAGTGCCGC
ProLeuAlaLeuAlaGlyProProProProProProProProHisProHisAlaArgIleLysLeuGluAsnProLeuAspTyrGlySerAlaTrpAlaAlaAlaAlaAlaGlnCysArg
TATGGGGACCTGGCGAGCCTGCATGGCGCGGGTGCAGCGGGACCCGGTTCTGGGTCACCCTCAGCCGCCGCTTCCTCATCCTGGCACACTCTCTTCACAGCCGAAGAAGGCCAGTTGTAT
TyrGlyAspLeuAlaSerLeuHisGlyAlaGlyAlaAlaGlyProGlySerGlySerProSerAlaAlaAlaSerSerSerTrpHisThrLeuPheThrAlaGluGluGlyGlnLeuTyr
GGACCGTGTGGTGGTGGTGGGGGTGGTGGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGAGGCGGAAGCTGTAGCCCCCTACGGCTAC
GlyProCysGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGluAlaGluAlaValAlaProTyrGlyTyr
ACTCGGCCCCCTCAGGGGCTGGCGGGCCAGGAAAGCGACTTCACCGCACCTGATGTGTGGTACCCTGGCGGCATGGTGAGCAGAGTGCCCTATCCCAGTCCCACTTGTGTCAAAAGCGAA
ThrArgProProGlnGlyLeuAlaGlyGlnGluSerAspPheThrAlaProAspValTrpTyrProGlyGlyMetValSerArgValProTyrProSerProThrCysValLysSerGlu
ATGGGCCCCTGGATGGATAGCTACTCCGGACCTTACGGGGACATGCGTTTGGAGACTGCCAGGGACCATGTTTTGCCCATTGACTATTACTTTCCACCCCAGAAGACCTGCCTGATCTGT
MetGlyProTrpMetAspSerTyrSerGlyProTyrGlyAspMetArgLeuGluThrAlaArgAspHisValLeuProIleAspTyrTyrPheProProGlnLysThrCysLeuIleCys
GGAGATGAAGCTTCTGGGTGTCACTATGGAGCT
GlyAspGluAlaSerGlyCysHisTyrGlyAla

CGC CCG GGG ATC CTC TAG
Arg Pro Gly Ile Leu STOP
```

5 amino acid linker

Total amino acid: 323 + 6 + 242 + 5 = 576

FIGURE 9B

```
    170         180         190         200         210         220         230         240         250
ATG CAA ACA CAA AAA CCG ACT CTC GAA CTG CTA ACC TGC GAA GGC GCT TAT CGC GAC AAT CCC ACC GCG CTT TTT CAC CAG TTG TGT GGG
MET GLN THR GLN LYS PRO THR LEU GLU LEU LEU THR CYS GLU GLY ALA TYR ARG ASP ASN PRO THR ALA LEU PHE HIS GLN LEU CYS GLY
 1                           10                            20                            30
    260         270         280         290         300         310         320         330         340
GAT CGT CCG GCA ACG CTG CTG CTG GAA TCC GCA GAT ATC GAC AGC AAA GAT GAT TTA AAA AGC CTG CTG CTG GTA GAC AGT GCG CTG CGC
ASP ARG PRO ALA THR LEU LEU LEU GLU SER ALA ASP ILE ASP SER LYS ASP ASP LEU LYS SER LEU LEU LEU VAL ASP SER ALA LEU ARG
                             40                            50                            60
    350         360         370         380         390         400         410         420         430
ATT ACA GCT TTA GGT GAC ACT GTC ACA ATC CAG GCA CTT TCC GGC AAC GGC GAA GCC CTC CTG GCA CTA CTG GAT AAC GCC CTG CCT GCG
ILE THR ALA LEU GLY ASP THR VAL THR ILE GLN ALA LEU SER GLY ASN GLY GLU ALA LEU LEU ALA LEU LEU ASP ASN ALA LEU PRO ALA
                             70                            80                            90
    440         450         460         470         480         490         500         510         520
GGT GTG GAA AGT GAA CAA TCA CCA AAC TGC CGT GTG CTG CGC TTC CCC CCT GTC AGT CCA CTG CTG GAT GAA GAC GCC CGC TTA TGC TCC
GLY VAL GLU SER GLU GLN SER PRO ASN CYS ARG VAL LEU ARG PHE PRO PRO VAL SER PRO LEU LEU ASP GLU ASP ALA ARG LEU CYS SER
                             100                           110                           120
    530         540         550         560         570         580         590         600         610
CTT TCG GTT TTT GAC GCT TTC CGT TTA TTG CAG AAT CTG TTG AAT GTA CCG AAG GAA GAA CGA GAA GCC ATG TTC TTC AGC GGC CTG TTC
LEU SER VAL PHE ASP ALA PHE ARG LEU LEU GLN ASN LEU LEU ASN VAL PRO LYS GLU GLU ARG GLU ALA MET PHE PHE SER GLY LEU PHE
                             130                           140                           150
    620         630         640         650         660         670         680         690         700
TCT TAT GAC CTT GTG GCG GGA TTT GAA GAT TTA CCG CAA CTG TCA GCG GAA AAT AAC TGC CCT GAT TTC TGT TTT TAT CTC GCT GAA ACG
SER TYR ASP LEU VAL ALA GLY PHE GLU ASP LEU PRO GLN LEU SER ALA GLU ASN ASN CYS PRO ASP PHE CYS PHE TYR LEU ALA GLU THR
                             160                           170                           180
    710         720         730         740         750         760         770         780         790
CTG ATG GTG ATT GAC CAT CAG AAA AAA AGC ACC CGT ATT CAG GCC AGC CTG TTT GCT CCG AAT GAA GAA GAA AAA CAA CGT CTC ACT GCT
LEU MET VAL ILE ASP HIS GLN LYS LYS SER THR ARG ILE GLN ALA SER LEU PHE ALA PRO ASN GLU GLU GLU LYS GLN ARG LEU THR ALA
                             190                           200                           210
    800         810         820         830         840         850         860         870         880
CGC CTG AAC GAA CTA CGT CAG CAA CTG ACC GAA GCC GCG CCG CCG CTG CCA GTG GTT TCC GTG CCG CAT ATG CGT TGT GAA TGT AAT CAG
ARG LEU ASN GLU LEU ARG GLN GLN LEU THR GLU ALA ALA PRO PRO LEU PRO VAL VAL SER VAL PRO HIS MET ARG CYS GLU CYS ASN GLN
                             220                           230                           240
    890         900         910         920         930         940         950         960         970
AGC GAT GAA GAG TTC GGT GGC GTA GTG CGT TTG TTG CAA AAA GCG ATT CGC GCT GGA GAA ATT TTC CAG GTG GTG CCA TCT CGC CGT TTC
SER ASP GLU GLU PHE GLY GLY VAL VAL ARG LEU LEU GLN LYS ALA ILE ARG ALA GLY GLU ILE PHE GLN VAL VAL PRO SER ARG ARG PHE
                             250                           260                           270
    980         990         1000        1010        1020        1030        1040        1050        1060
TCT CTG CCC TGC CCG TCA CCG CTG GCG GCC TAT TAC GTG CTG AAA AAG AGT AAT CCC AGC CCG TAC ATG TTT TTT ATG CAG GAT AAT GAT
SER LEU PRO CYS PRO SER PRO LEU ALA ALA TYR TYR VAL LEU LYS LYS SER ASN PRO SER PRO TYR MET PHE PHE MET GLN ASP ASN ASP
                             280                           290                           300
    1070        1080        1090        1100        1110        1120        1130
TTC ACC CTA TTT GGC GCG TCG CCG GAA AGC TCG CTC AAG TAT GAT GCC ACC AGC CGC CAG ATT GAG ATC
PHE THR LEU PHE GLY ALA SER PRO GLU SER SER LEU LYS TYR ASP ALA THR SER ARG GLN ILE GLU ILE
                             310                           320
```

11 amino acid linker

```
CCC GGG CGA GCT CGA ATT CGA GCT CGC CCG GGG
Pro Gly Arg Ala Arg Ile Arg Ala Arg Pro Gly
```

FIGURE 10A

```
                                                       ATC TGT GGA GAT GAA GCT TCT GGT TGT CAC TAC GGA GCT CTC ACT TGT GGC AGC
                                                       ile cys gly asp glu ala ser gly cys his tyr gly ala leu thr cys gly ser 561
TGC AAG GTC TTC TTC AAA AGA GCT GCG GAA GGG AAA CAG AAG TAT CTA TGT GCC AGC AGA AAT GAT TGC ACC ATT GAT AAA TTT CGG AGG
cys lys val phe phe lys arg ala ala glu gly lys gln lys tyr leu cys ala ser arg asn asp cys thr ile asp lys phe arg arg 591
AAA AAT TGT CCA TCG TGT CGT CTC CGG AAA TGT TAT GAA GCA GGG ATG ACT CTG GGA GCT CGT AAG CTG AAG AAA CTT GGA AAT CTC AAA
lys asn cys pro ser cys arg leu arg lys cys tyr glu ala gly met thr leu gly ala arg lys leu lys lys leu gly asn leu lys 621
CTA CAG GAA GAA GGA GAA AAC TCC AGT GCT GGT AGC CCC ACT GAG GAC CCA TCC CAG AAG ATG ACT GTA TCA CAC ATT GAA GGC TAT GAA
leu gln glu glu gly glu asn ser ser ala gly ser pro thr glu asp pro ser gln lys met thr val ser his ile glu gly tyr glu 651
TGT CAA CCT ATC TTT CTT AAT GTC CTG GAA GCC ATT GAG CCA GGA GTG GTG TGT GCC GGA CAT GAC AAC AAC CAG CCT GAT TCC TTT GCT
cys gln pro ile phe leu asn val leu glu ala ile glu pro gly val val cys ala gly his asp asn asn gln pro asp ser phe ala 681
GCC TTG TTA TCT AGT CTC AAC GAG CTT GGC GAG AGA CAG CTT GTA CAT GTG GTC AAG TGG GCC AAG GCC TTG CCT GGC TTC CGC AAC TTG
ala leu leu ser ser leu asn glu leu gly glu arg gln leu val his val val lys trp ala lys ala leu pro gly phe arg asn leu 711
CAT GTG GAT GAC CAG ATG GCA GTC ATT CAG TAT TCC TGG ATG GGA CTG ATG GTA TTT GCC ATG GGT TGG CGG TCC TTC ACT AAT GTC AAC
his val asp asp gln met ala val ile gln tyr ser trp met gly leu met val phe ala met gly trp arg ser phe thr asn val asn 741
TCT AGG ATG CTC TAC TTT GCA CCT GAC CTG GTT TTC AAT GAG TAT CGC ATG CAC AAG TCT CGA ATG TAC AGC CAG TGC GTG AGG ATG AGG
ser arg met leu tyr phe ala pro asp leu val phe asn glu tyr arg met his lys ser arg met tyr ser gln cys val arg met arg 771
CAC CTT TCT CAA GAG TTT GGA TGG CTC CAG ATA ACC CCC CAG GAA TTC CTG TGC ATG AAA GCA CTG CTA CTC TTC AGC ATT ATT CCA GTG
his leu ser gln glu phe gly trp leu gln ile thr pro gln glu phe leu cys met lys ala leu leu leu phe ser ile ile pro val 801
GAT GGG CTG AAA AAT CAA AAA TTC TTT GAT GAA CTT CGA ATG AAC TAC ATC AAG GAA CTT GAT
asp gly leu lys asn gln lys phe phe asp glu leu arg met asn tyr ile lys glu leu asp

CCT CTA GAG TCG ACC TGC AGC CCA AGC TTA TCG ATG ATA AGC TGT CAA ACA TGA
Pro Leu Glu Ser Thr Cys Ser Pro Ser Leu Ser Met Ile Ser Cys Gln Thr ATOP
                                      17 amino acid linker
```

Total amino acids: 323 + 11 + 279 + 17 = 630

FIGURE 10B

323 amino acids from TRP E protein

```
     170         180         190         200         210         220         230         240         250
ATG CAA ACA CAA AAA CCG ACT CTC GAA CTG CTA ACC TGC GAA GGC GCT TAT CGC GAC AAT CCC ACC GCG CTT TTT CAC CAG TTG TGT GGG
MET GLN THR GLN LYS PRO THR LEU GLU LEU LEU THR CYS GLU GLY ALA TYR ARG ASP ASN PRO THR ALA LEU PHE HIS GLN LEU CYS GLY
 1                                   10                                      20                                      30
     260         270         280         290         300         310         320         330         340
GAT CGT CCG GCA ACG CTG CTG CTG GAA TCC GCA GAT ATC GAC AGC AAA GAT GAT TTA AAA AGC CTG CTG CTG GTA GAC AGT GCG CTG CGC
ASP ARG PRO ALA THR LEU LEU LEU GLU SER ALA ASP ILE ASP SER LYS ASP ASP LEU LYS SER LEU LEU LEU VAL ASP SER ALA LEU ARG
                                     40                                      50                                      60
     350         360         370         380         390         400         410         420         430
ATT ACA GCT TTA GGT GAC ACT GTC ACA ATC CAG GCA CTT TCC GGC AAC GGC GAA GCC CTC CTG GCA CTA CTG GAT AAC GCC CTG CCT GCG
ILE THR ALA LEU GLY ASP THR VAL THR ILE GLN ALA LEU SER GLY ASN GLY GLU ALA LEU LEU ALA LEU LEU ASP ASN ALA LEU PRO ALA
                                     70                                      80                                      90
     440         450         460         470         480         490         500         510         520
GGT GTG GAA AGT GAA CAA TCA CCA AAC TGC CGT GTG CTG CGC TTC CCC CCT GTC AGT CCA CTG CTG GAT GAA GAC GCC CGC TTA TGC TCC
GLY VAL GLU SER GLU GLN SER PRO ASN CYS ARG VAL LEU ARG PHE PRO PRO VAL SER PRO LEU LEU ASP GLU ASP ALA ARG LEU CYS SER
                                    100                                     110                                     120
     530         540         550         560         570         580         590         600         610
CTT TCG GTT TTT GAC GCT TTC CGT TTA TTG CAG AAT CTG TTG AAT GTA CCG AAG GAA GAA CGA GAA GCC ATG TTC TTC AGC GGC CTG TTC
LEU SER VAL PHE ASP ALA PHE ARG LEU LEU GLN ASN LEU LEU ASN VAL PRO LYS GLU GLU ARG GLU ALA MET PHE PHE SER GLY LEU PHE
                                    130                                     140                                     150
     620         630         640         650         660         670         680         690         700
TCT TAT GAC CTT GTG GCG GGA TTT GAA GAT TTA CCG CAA CTG TCA GCG GAA AAT AAC TGC CCT GAT TTC TGT TTT TAT CTC GCT GAA ACG
SER TYR ASP LEU VAL ALA GLY PHE GLU ASP LEU PRO GLN LEU SER ALA GLU ASN ASN CYS PRO ASP PHE CYS PHE TYR LEU ALA GLU THR
                                    160                                     170                                     180
     710         720         730         740         750         760         770         780         790
CTG ATG GTG ATT GAC CAT CAG AAA AAA AGC ACC CGT ATT CAG GCC AGC CTG TTT GCT CCG AAT GAA GAA GAA AAA CAA CGT CTC ACT GCT
LEU MET VAL ILE ASP HIS GLN LYS LYS SER THR ARG ILE GLN ALA SER LEU PHE ALA PRO ASN GLU GLU GLU LYS GLN ARG LEU THR ALA
                                    190                                     200                                     210
     800         810         820         830         840         850         860         870         880
CGC CTG AAC GAA CTA CGT CAG CAA CTG ACC GAA GCC GCG CCG CCG CTG CCA GTG GTT TCC GTG CCG CAT ATG CGT TGT GAA TGT AAT CAG
ARG LEU ASN GLU LEU ARG GLN GLN LEU THR GLU ALA ALA PRO PRO LEU PRO VAL VAL SER VAL PRO HIS MET ARG CYS GLU CYS ASN GLN
                                    220                                     230                                     240
     890         900         910         920         930         940         950         960         970
AGC GAT GAA GAG TTC GGT GGC GTA GTG CGT TTG TTG CAA AAA GCG ATT CGC GCT GGA GAA ATT TTC CAG GTG GTG CCA TCT CGC CGT TTC
SER ASP GLU GLU PHE GLY GLY VAL VAL ARG LEU LEU GLN LYS ALA ILE ARG ALA GLY GLU ILE PHE GLN VAL VAL PRO SER ARG ARG PHE
                                    250                                     260                                     270
     980         990        1000        1010        1020        1030        1040        1050        1060
TCT CTG CCC TGC CCG TCA CCG CTG GCG GCC TAT TAC GTG CTG AAA AAG AGT AAT CCC AGC CCG TAC ATG TTT TTT ATG CAG GAT AAT GAT
SER LEU PRO CYS PRO SER PRO LEU ALA ALA TYR TYR VAL LEU LYS LYS SER ASN PRO SER PRO TYR MET PHE PHE MET GLN ASP ASN ASP
                                    280                                     290                                     300
    1070        1080        1090        1100        1110        1120        1130
TTC ACC CTA TTT GGC GCG TCG CCG GAA AGC TCG CTC AAG TAT GAT GCC ACC AGC CGC CAG ATT GAG ATC
PHE THR LEU PHE GLY ALA SER PRO GLU SER SER LEU LYS TYR ASP ALA THR SER ARG GLN ILE GLU ILE
                                    310                                     320
```

CCC CCG
Pro Pro

FIGURE 11A

2 amino acid linker

```
                                                            GAA TTC CTG TGC ATG AAA GCA CTG CTA CTC TTC AGC ATT ATT CCA GTG
                                                            glu phe leu cys met lys ala leu leu leu phe ser ile ile pro val 401
GAT GGG CTG AAA AAT CAA AAA TTC TTT GAT GAA CTT CGA ATG AAC TAC ATC AAG GAA CTT GAT CGC ATC ATT GCA TGC AAA AGA AAA AAT
asp gly leu lys asn gln lys phe phe asp glu leu arg met asn tyr ile lys glu leu asp arg ile ile ala cys lys arg lys asn 431
CCC ACA TCC TGC TCA AGG CGC TTC TAC CAG CTC ACC AAG CTC CTG GAT TCT GTG CAG CCT ATT GCA AGA GAG CTG CAT CAA TTC ACT TTT
pro thr ser cys ser arg arg phe tyr gln leu thr lys leu leu asp ser val gln pro ile ala arg glu leu his gln phe thr phe 461
GAC CTG CTA ATC AAG TCC CAT ATG GTG AGC GTG GAC TTT CCT GAA ATG ATG GCA GAG ATC ATC TCT GTG CAA GTG CCC AAG ATC CTT TCT
asp leu leu ile lys ser his met val ser val asp phe pro glu met met ala glu ile ile ser val gln val pro lys ile leu ser 491
GGG AAA GTC AAG CCC ATC TAT TTC CAC ACA CAG TGA
gly lys val lys pro ile tyr phe his thr gln ***
```

Total amino acids: 323 + 2 + 117 = 442

FIGURE 11B

INTERNATIONAL SEARCH REPORT

International Application No PCT/US90/06015

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all) [3]

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC(5): C07H 21/00; C12N 1/20; C12N 15/00; C07K 13/00

U.S. CL.: 536/27; 435/6,7,240.2, 252.3, 317.1; 530/350,387

II. FIELDS SEARCHED

Minimum Documentation Searched [4]

Classification System	Classification Symbols
U.S.	536/27; 530/350,387; 435/69.1, 172.3, 240.2, 252.3, 317.1, 6, 7; 935/6,22,27,32,70,111

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched [5]

Databases: DIALOG (Files 55,311,312,154), USPTO Automated Patent System (File USPAT, 1971-1990). See attachment for search terms.

III. DOCUMENTS CONSIDERED TO BE RELEVANT [14]

Category *	Citation of Document, [16] with indication, where appropriate, of the relevant passages [17]	Relevant to Claim No. [18]
X / Y	WO, A, 89/09791 (FRENCH ET AL.), 19 October 1989, see the entire document.	1-3,5,16,18,19 / 4,6,7,17,40,42
X	WO, A, 89/09223 (LIAO ET AL.), 05 October 1989, see the entire document.	1-7,16-19,40, 42
A	Science, Volume 240, Issued 13 May 1988 (Washington, USA), Evans, "The Steroid and thyroid hormone receptor superfamily", pages 889-895, see the entire document.	
X / Y	Progress in Cancer Research and Therapy, (Raven Press, New York, USA) Volume 35, Issued July 1988, Govindan et al., "Cloning of the human androgen receptor cDNA", pages 49-54, see the entire document.	1,2,5,16,18,19 / 3,4,6,7,17,40, 42

* Special categories of cited documents: [15]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search [2]	Date of Mailing of this International Search Report [2]
16 January 1991	12 FEB 1991
International Searching Authority [1]	**Signature of Authorized Officer** [10]
ISA/US	Jasemine C. Chambers Jasemine C. Chambers ebw

III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)

Category *	Citation of Document, [16] with indication, where appropriate, of the relevant passages [17]	Relevant to Claim No [18]
X Y	Biochemical and Biophysical Research Communications, Volume 153, no. 1, Issued 31 May 1988, (Academic Press, Orlando, USA) Trapman et al., "Cloning, structure and expression of a cDNA encoding the human androgen receptor", pages 241-248, see entire document.	1,2,5,16,18,19 3,4,7,17,40,42,6
X Y	Proceedings of the National Academy of Sciences, Volume 85, Issued October 1988 (Washington, USA) Chang et al., "Structural analysis of complementary DNA and amino acid sequences of human and rat androgen receptors", pages 7211-7215, see the entire document.	1-5,16,18,19 3, 6,7,17,40,42
X Y	Science, Volume 240, Issued 15 April 1988, (Washington, USA), Lubahn et al., "Cloning of human androgen receptor complementary DNA and localization to the X chromosome", pages 327-330, see the entire document.	1,2,5,6,16,18,19, 3,4,7,17,40,42
X Y	Science, Volume 240, Issued 15 April 1988, (Washington USA), Chang et al., "Molecular cloning of human and rat complementary DNA encoding androgen receptor", pages 324-326, see the entire document.	1-3,5,16,18,19 4,6,7,17,40,42
X Y	Proceedings of the National Academy of Sciences, Volume 86, Issued January 1989 (Washington, USA), Tilley et al., "Characterization and expression of a cDNA encoding the human androgen receptor", pages 327-331, see the entire document.	1,2,5,6,16,18,19 3,4,6,7,17,40,42
X Y	Molecular Endocrinology, Volume 2, Number 12, Issued December 1988 (Baltimore, USA), Lubahn et al., "The human androgen receptor: complementary deoxyribonucleic acid cloning, sequence analysis and gene expression in prostate", pages 1265-1275, see the entire document.	1,2,5,6,16,18,19 3,4,7,17,40,42

III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)

Category *	Citation of Document, [16] with indication, where appropriate, of the relevant passages [17]	Relevant to Claim No [18]
X Y	Molecular Endocrinology, Volume 2, Number 12, Issued December 1988 (Baltimore, USA), Tan et al., "The rat androgen receptor: primary structure, autoregulation of its messenger ribonucleic acid, and immunocytochemical localization of the receptor protein", pages 1276-1285, see the entire document.	1,3,5,16,18,19 2,4,6,7,17,40,42
X,P Y	Proceedings of the National Academy of Sciences, Volume 86, Issued December 1989 (Washington, USA), Lubahn et al., "Sequence of the intron/exon junctions of the coding region of the human androgen receptor gene and identification of a point mutation in a family with complete androgen insensitivity", pages 9534-9538, see the entire document.	1,2,6,16,18,40,42 3,4,5,7,17,19
Y	Cold Spring Harbor Symposia on Quantitative Biology, Volume LI, Published 1986, (Cold Spring Harbor Laboratory, New York, USA), Mullis et al., "Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction", pages 263-273, see the entire document.	40,42
Y	Nature, Volume 324, Issued 13 November 1986, (London, UK) Saiki et al., "Analysis of enzymatically amplified B-globin and HLA-DQ DNA with allele-specific oligonucleotide probes", pages 163-166, see the entire document.	40,42
Y	US, A, 4,800,159 (MULLIS et al.) 24 January 1989, see the entire document.	40,42

FURTHER INFORMATION CONTINUED FROM THE SECOND SHEET

V. ☐ OBSERVATIONS WHERE CERTAIN CLAIMS WERE FOUND UNSEARCHABLE [1]

This international search report has not been established in respect of certain claims under Article 17(2) (a) for the following reasons:

1. ☐ Claim numbers , because they relate to subject matter [1] not required to be searched by this Authority, namely:

2. ☐ Claim numbers ...., because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out [1], specifically:

3. ☐ Claim numbers__________, because they are dependent claims not drafted in accordance with the second and third sentences of PCT Rule 6.4(a).

VI. ☒ OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING [2]

This International Searching Authority found multiple inventions in this international application as follows:

See attachment.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims of the international application.

2. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims of the international application for which fees were paid, specifically claims:

3. ☒ No required additional search fees were timely paid by the applicant. Consequently, this International search report is restricted to the invention first mentioned in the claims; it is covered by claim numbers:

   1-7,16-19,40,42 (telephone practice)

4. ☐ As all searchable claims could be searched without effort justifying an additional fee, the International Searching Authority did not invite payment of any additional fee.

Remark on Protest

☐ The additional search fees were accompanied by applicant's protest.

☐ No protest accompanied the payment of additional search fees.

Attachment to PCT/ISA/210

Part VI. Observations where unity of invention is lacking

I. Claims 1-7, 16-19, 40 and 42, drawn to androgen receptor DNA, plasmid, cell and a method of use of the DNA, Class 435, subclasses 6, 240.2, 252.3 and 317.1, and Class 536, subclass 27.

II. Claims 8-15, 16-19, 41 and 42, drawn to TR2 DNA, plasmid, cell and a method of use of the DNA, Class 435, subclasses 6, 240.2, 252.3 and 317.1, and Class 536, subclass 27.

III. Claims 20, 21, 24-26 and 33, drawn to androgen receptor polypeptides and a method of making the same, Classes 530 and 435, subclasses 350 and 69.1, respectively.

IV. Claims 22, 23, 27-32 and 33, drawn to TR2 polypeptides and a method of making the same, Classes 530 and 435, subclasses 350 and 69.1, respectively.

V. Claims 34-38, drawn to an antibody reactive with androgen receptors and a method of using the same, Classes 530 and 435, subclasses 387 and 7, respectively.

VI. Claims 34-37 and 39, drawn to an antibody reactive with TR2 polypeptides and a method of using the same, Classes 530 and 435, subclasses 387 and 7, respectively.

Attachment to PCT/ISA/210,
Part II.

II. FIELDS SEARCHED SEARCH TERMS:

androgen receptor, human, rat, gene, sequence, cDNA cloning, express, hybridization, review, inventors' names.